
NWPNLEH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 12 11:49:45 1998; MasPar time 24.12 seconds
Tabular output not generated.

Title: >US-08-083-590A-2
Description: (1-833) from US08083590A.pep (1 of 2)
Perfect Score: 6164
Sequence: 1 MHMVKLLTAFICFIVIVQV.....GSGAGTAQQQSRVCGTPEHM 833

Scoring table:
PAM 150
Gap 11

Searched: 124785 segs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseg31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 36.728; Variance 180.698; scale 0.203

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-----------|-----------------------|-----------|
| 1 | 6143 | 99.7 | 833 | 6 R28960 | Delta D11 | 0.00e+00 |
| 2 | 1910 | 31.0 | 727 | 21 W11719 | C-Delta-1 polypeptide | 9.35e-151 |
| 3 | 1910 | 31.0 | 727 | 21 W00876 | C-Delta-1 polypeptide | 9.35e-151 |
| 4 | 1861 | 30.2 | 723 | 25 W18353 | Proliferation and dif | 1.78e-146 |
| 5 | 1839 | 29.8 | 723 | 25 W18348 | Proliferation and dif | 1.44e-144 |
| 6 | 1839 | 29.8 | 702 | 25 W18349 | Proliferation and dif | 1.44e-144 |
| 7 | 1826 | 29.6 | 722 | 21 W11720 | M-Delta-1 polypeptide | 1.94e-143 |
| 8 | 1784 | 28.9 | 726 | 7 R38305 | Sequence of a delta p | 8.50e-140 |
| 9 | 1718 | 27.9 | 1193 | 19 W05835 | Chick Serrate-1 | 4.45e-134 |
| 10 | 1692 | 27.4 | 1218 | 19 W05833 | Human Serrate-1 (Hn) | 7.98e-132 |
| 11 | 1680 | 27.3 | 1218 | 25 W18354 | Proliferation and dif | 8.74e-131 |
| 12 | 1676 | 27.2 | 1036 | 25 W18351 | Proliferation and dif | 1.94e-130 |
| 13 | 1676 | 27.2 | 1187 | 25 W18352 | Proliferation and dif | 1.94e-130 |
| 14 | 1346 | 21.8 | 1257 | 19 W05834 | Human Serrate-2 (Hn) | 6.38e-102 |
| 15 | 1258 | 20.4 | 1404 | 7 R38304 | Sequence of a serrate | 2.42e-94 |
| 16 | 883 | 14.3 | 660 | 21 W11726 | H-Delta-1 polypeptide | 3.14e-62 |
| 17 | 878 | 14.2 | 192 | 21 W11728 | H-Delta-1 polypeptide | 8.36e-62 |
| 18 | 727 | 11.8 | 200 | 25 W18347 | Proliferation and dif | 5.31e-49 |
| 19 | 686 | 11.1 | 473 | 17 R86869 | Adhesive protein. | 1.52e-45 |

| | | | | | | |
|----|-----|------|------|-----------|-----------------------|----------|
| 20 | 660 | 10.7 | 383 | 10 R56166 | Neuroendocrine tumor | 2.34e-43 |
| 21 | 654 | 10.6 | 385 | 10 R56167 | Neuroendocrine tumor | 7.48e-43 |
| 22 | 646 | 10.5 | 1480 | 5 R25079 | Drosophila SLIT prote | 3.51e-42 |
| 23 | 610 | 9.9 | 660 | 21 W11725 | H-Delta-1 polypeptide | 3.66e-39 |
| 24 | 551 | 8.9 | 157 | 21 W11730 | H-Delta-1 polypeptide | 3.08e-34 |
| 25 | 544 | 8.8 | 192 | 25 W18350 | Proliferation and dif | 1.18e-33 |
| 26 | 516 | 8.4 | 293 | 6 R28961 | Serrate protein. | 2.49e-31 |
| 27 | 493 | 8.0 | 2199 | 16 R45652 | Human cytotactin. | 1.99e-29 |
| 28 | 470 | 7.6 | 1810 | 16 R45653 | Chicken cytotactin. | 1.57e-27 |
| 29 | 446 | 7.2 | 2189 | 1 R05222 | Antigen G83401FL enco | 1.48e-25 |
| 30 | 388 | 6.3 | 1251 | 14 R79475 | Mouse LTB-3. | 8.01e-21 |
| 31 | 388 | 6.3 | 1833 | 14 R79478 | Mouse LTB-2. | 8.01e-21 |
| 32 | 379 | 6.1 | 753 | 10 R53088 | Human masking protein | 4.30e-20 |
| 33 | 379 | 6.1 | 752 | 10 R53087 | Human masking protein | 4.30e-20 |
| 34 | 379 | 6.1 | 756 | 10 R53086 | Human masking protein | 4.30e-20 |
| 35 | 379 | 6.1 | 845 | 10 R53085 | Human masking protein | 4.30e-20 |
| 36 | 369 | 6.0 | 1355 | 3 R24584 | TGF beta 1 binding pr | 2.77e-19 |
| 37 | 364 | 5.9 | 1712 | 4 R22461 | Masking protein high | 7.01e-19 |
| 38 | 346 | 5.6 | 810 | 27 W37500 | Human nrl-related pro | 1.97e-17 |
| 39 | 348 | 5.6 | 816 | 27 W37501 | Human nrl-related pro | 1.36e-17 |
| 40 | 307 | 5.0 | 2707 | 24 W27161 | Mouse receptor M22. | 2.55e-14 |
| 41 | 287 | 4.7 | 77 | 6 R28962 | ELR-11 and -12. | 9.62e-13 |
| 42 | 290 | 4.7 | 1257 | 9 R46627 | Neurocan core protein | 5.62e-13 |
| 43 | 292 | 4.7 | 2409 | 3 R12609 | Version. | 3.91e-13 |
| 44 | 291 | 4.7 | 4544 | 9 R47861 | Alpha 2-Macroglobulin | 4.69e-13 |
| 45 | 291 | 4.7 | 4544 | 11 R60517 | Human alpha-2-MR. | 4.69e-13 |

ALIGNMENTS

RESULT 1
ID R28960 standard; Protein; 833 AA.

AC R28960; 01-APR-1993 (first entry)
DE Delta D11.
KW Human; Notch; plasmid; cDNA; clone; D11; expression library; PCR;
KW Humanase chain reaction; primer; cloning vector; Delta; Serrate;
KW neurogenic; topolythmic; homotypic; heterotypic; differentiation;
KW quantitation; antibody.
OS Homo sapiens.
PN W09219734-A.
PD 12-NOV-1992.
PE 01-MAY-1992; U03651.
PR 03-MAY-1991; US-695189.
PA 14-NOV-1991; US-791923.
PA (INDV) UNIV INDIANA FOUND.
PI (UYVA) UNIV YALE.
PI Artavanis-Tsakonas S, Blaumueller CM, Fehon RG, Muskavitch MAT;
PI Rebay I, Shepard SB;
DR WPI; 92-398861/48.
DR N-PSDB; Q30997.
PT Human Notch and Delta DNA and protein sequences - used for study
PT and manipulation of differentiation processes
PS Claim 50; Fig 13; 23pp; English.
CC The sequence given is encoded by the nucleotide sequence of human
CC Delta gene contained in plasmid cDNA clone D11. A human expression
CC library was constructed and screening assays were carried out on to
CC select for the expressed Delta product. Alternatively the sequences
CC could be isolated by amplification using polymerase chain reaction
CC (PCR) primers. The isolated gene may be inserted into a cloning
CC vector and expressed. The Delta gene and also the Notch and Serrate
CC neurogenic genes are designated "topolythmic" genes. The proteins
CC they encode are involved in specific homo- or heterotypic interactions
CC crucial to differentiation. The quantitation of mRNA for human Notch
CC and Delta and adhesive molecules, and study of its expression are
CC possible using the DNA and antibodies raised against the Notch and
CC Delta proteins.
SQ Sequence 833 AA;

Query Match 99.7%; Score 6143; DB 6; Length 833;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 831; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Db 1 mhwikclltafcivvqshssqfelrlkylfnsdhrdnegrcssesdgaatgkclgs 60
Qy 1 mhwikclltafcivvqshssqfelrlkylfnsdhrdnegrcssesdgaatgkclgs 60
Db 61 ktrfrvclkyqatidttstqctygdvtpilgensvnltdagrfnkgftnpiqpfss 120
Qy 61 ktrfrvclkyqatidttstqctygdvtpilgensvnltdagrfnkgftnpiqpfss 120
Db 121 swpftsliveawhdttnsgnartnklllgrllvqvqvlevsswtknksesqytsleydf 180
Qy 121 swpftsliveawhdttnsgnartnklllgrllvqvqvlevsswtknksesqytsleydf 180
Db 181 rvtcdlnvygscakfcrprddsfghstcsetgeiicltgwgqdychpkcakgcehgbc 240
Qy 181 rvtcdlnvygscakfcrprddsfghstcsetgeiicltgwgqdychpkcakgcehgbc 240
Db 241 dkrncvcqlgwkalcnecvlepnchgtcnkpwcticnecwgglycnqdlmynctnhrp 300
Qy 241 dkrncvcqlgwkalcnecvlepnchgtcnkpwcticnecwgglycnqdlmynctnhrp 300
Db 301 knggtcftatgclircpangtclnlnndyqctcragftgkdsviddecssgpcn 540
Qy 301 knggtcftatgclircpangtclnlnndyqctcragftgkdsviddecssgpcn 540
Db 361 hcangwqkmeekvltcsdkpchgicrnvrpqlgskqgqyccpipygqncdlql 420
Qy 361 hcangwqkmeekvltcsdkpchgicrnvrpqlgskqgqyccpipygqncdlql 420
Db 421 ncsnpncingscqpsgkicpagfsagtrcctnidclghqcnngtctdmvnyqrcqv 480
Qy 421 ncsnpncingscqpsgkicpagfsagtrcctnidclghqcnngtctdmvnyqrcqv 480
Db 481 pfghthesskvdclircpangtclnlnndyqctcragftgkdsviddecssgpcn 540
Qy 481 pfghthesskvdclircpangtclnlnndyqctcragftgkdsviddecssgpcn 540
Db 541 gqtcmmrvnsfecnangfrgkqcdesvdfdahyqatqaradgltnaqqvllav 600
Qy 541 gqtcmmrvnsfecnangfrgkqcdesvdfdahyqatqaradgltnaqqvllav 600
Db 601 fsavmplvaviacvfmckrkrkraqekdaearkqneqnavatmnhngsgvgvalasa 660
Qy 601 fsavmplvaviacvfmckrkrkraqekdaearkqneqnavatmnhngsgvgvalasa 660
Db 661 slgkgtsgnslgtfdggnpnlkntwksvnnicasaasaaadecclmyggvvasv 720
Qy 661 slgkgtsgnslgtfdggnpnlkntwksvnnicasaasaaadecclmyggvvasv 720
Db 721 adnnnansvfcvaplqraksqkqlntdptlmhrqspagsakagsggggaaqkrisvl 780
Qy 721 adnnnansvfcvaplqraksqkqlntdptlmhrqspagsakagsggggaaqkrisvl 780
Db 781 gqgyscqrwpslaaagvagacssqlmaasaaagsgagtaqqrvsvvcgtpm 833
Qy 781 gqgyscqrwpslaaagvagacssqlmaasaaagsgagtaqqrvsvvcgtpm 833

RESULT 2
ID W11719 standard; Protein: 727 AA.
AC W11719;
DT 28-APR-1997 (first entry)
DE C-Delta-1 polypeptide.
KW C-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;
KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy.
OS Gallus sp.
FH key Location/Qualifiers
FT domain 184..228
FT /label= DSL
FT domain 229..261
FT /label= EGF1
FT domain 262..292
```

```
FT domain /label= EGF2
FT 293..332
FT /label= EGF3
FT 333..370
FT /label= EGF4
FT 371..409
FT /label= EGF5
FT 410..447
FT /label= EGF6
FT 448..485
FT /label= EGF7
FT 486..523
FT /label= EGF8
FT 524..534
FT /label= EGF9
FT 555..579
FT /label= TM
FT /note= "transmembrane domain"
PN WO9701571-A1.
PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYVA ) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
PI Lewis J;
PI WPI; 97-100159/09.
DR N-PSDB; T58897.
DT New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PS Disclosure; Fig 2: 135pp; English.
CC C-delta-1 polypeptide (W11719) is the chick homologue of Drosophila
CC Delta, a protein that binds to Notch protein. Expression of
CC C-Delta-1 correlates with onset of neurogenesis. The C-delta-1
CC amino acid sequence was deduced from a cDNA clone (T58897) obtd.
CC from chick stage 4-6 embryos. An alternatively spliced variant
CC (W00876) was also isolated, and mouse (W11720) and human (W11721-
CC 38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC can be used to treat or prevent disorders characterised by
CC increased Notch activity, such as cervical, breast, lung or colon
CC cancer, melanoma or seminoma, and nervous system disorders or to
CC promote tissue regeneration and repair.
SQ Sequence 727 AA;
Query Match 31.0%; Score 1910; DB 21; Length 727;
Best Local Similarity 50.1%; Pred. No. 9,95e-151;
Matches 263; Conservative 78; Mismatches 168; Indels 16; Gaps 15;
Db 20 qvdsgsvfelklqefvnnkglslrnccrsgggpggagggqcdctkfrvclkyqavsp 79
Qy 19 QVHSSGFELRLKYFSNDHRDNegrcssesdgaatgkclgsktrfrvclkyqatid 78
Db 80 epctgysaitpvlansfsvpdgagadpafsnprfpgftwptfsliealhtdsp 139
Qy 79 TSQCTTGDVITPILGENSVNLTDAQRFPQNGFTNPQFPFSWPGFTSLIVCAWH-DTN 137
Db 140 ddlttenperlisrlatqrhlavgeewsqdlhssgrtdlkyysrfvdehygecsvfc 199
Qy 138 NSGNARTNKLLIQRLVQVLEVSSSEWTKNKSQYTSLEYDFRVTCDLNLYSGCAKFC 197
Db 200 rpddrfghfctgcekvncpnpwgqyctepiclpqcdqhgfdckpgeckrvvgqr 259
Qy 198 RPDDSFHSTCSETGEIICLTGQWGDYCHPKCAKGC-E-HGCDKPNQCVQLGKGA 255
Db 260 ydcirypcblhgtcqqpwcqgcqwgglfcnqdlmyncthhkpcnkatctntggys 319
Qy 256 LCNECVLEPNCHGTCKNKPWCTICNEGWGGLYCNQDLNLTNHRPCKNGTCTNTGG 315
Db 320 tscsrpgygtssceieinecda--npcnngsgctd-le-ns-ysctcpypfygkncelsa 374
Qy 316 TKCAPGYSDDCENEIYSCDADVNPQNGTCTCIDEPHTTKYKCHCANGSGKCEKV 375
```

| | | | |
|--------|---------------------------------------------------------------------|--------------------------------------------------------------|-----|
| Db | 375 | mtcagdgpfmg-r-c-td-npdy-gyscrebplysgfncexk1idyesspncangagcvd | 429 |
| Oy | 376 | LTCSKPKHQGICRNVRPGLSLKGGCYGCECPICGSGPNCDLQJDNCSNFCINIGGSC-Q | 434 |
| Db | 430 | lgnyslccgagaftrhcdhdvddcaasfpcvngtfcqdyvndysctoppysngkncstpv | 489 |
| Oy | 435 | -PSG-KCICPAGFSGSTRRETINIDDLGHCENGCGCICDMVQYKCGCVPEFGRHCSKV | 492 |
| Db | 490 | srcehncpchgatcbersnryvceccaryg9lncqfllpeppg9 | 534 |
| Oy | 493 | DLCLIRPCANGGTCLNNNDYQCTCRAGFTGKDCSDVIDECCSGP | 537 |
| RESULT | 3 | | |
| ID | W00876 | standard; Protein; 740 AA. | |
| AC | W00876; | | |
| DR | 28-APR-1997 | (first entry) | |
| DE | C-delta-1 polypeptide (alternatively spliced variant). | | |
| KM | C-delta-1, cell proliferation; nervous system disorder; | | |
| KM | tissue regeneration; Notch; cervix cancer; breast cancer; | | |
| KM | lung cancer; colon cancer; melanoma; seminoma; | | |
| KW | neurogenesis; therapy. | | |
| OS | Gallus sp. | | |
| FH | Key | location/qualifiers | |
| FT | domain | 184..228 | |
| FT | FT | /label- DSL | |
| FT | FT | 229..261 | |
| FT | FT | /label- EGF1 | |
| FT | FT | 262..292 | |
| FT | FT | /label- EGF2 | |
| FT | FT | 293..332 | |
| FT | FT | /label- EGF3 | |
| FT | FT | 333..370 | |
| FT | FT | /label- EGF4 | |
| FT | FT | 371..409 | |
| FT | FT | /label- EGF5 | |
| FT | FT | 410..447 | |
| FT | FT | /label- EGF6 | |
| FT | FT | 448..485 | |
| FT | FT | /label- EGF7 | |
| FT | FT | 486..523 | |
| FT | FT | /label- EGF8 | |
| FT | FT | 524..534 | |
| FT | FT | /label- EGF9 | |
| FT | FT | 535..579 | |
| FT | FT | /label- TM | |
| FT | FT | /note= "transmembrane domain" | |
| PN | W09701571-A1. | | |
| PD | 16-JAN-1997. | | |
| PF | 28-JUN-1996; U11178. | | |
| PR | 28-JUN-1995; US-000589. | | |
| PA | (IMCR) IMPERIAL CANCER RES TECHNOLOGY. | | |
| PA | (UYVA) UNIV YALE. | | |
| PI | Artevanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D; | | |
| PI | Lewis J; | | |
| DR | WPI; 97-100159/09. | | |
| DR | N-PSDB; T58898. | | |
| PT | New vertebrate Delta protein, DNA and antibodies - for treating and | | |
| PT | preventing cancer, nervous system disorders and for tissue | | |
| PT | regeneration | | |
| PS | Disclosure; Fig 2; 135pp; English. | | |
| CC | C-delta-1 polypeptide (W00876) is the chick homologue of Drosophila | | |
| CC | Delta, a protein that binds to Notch protein. Expression of | | |
| CC | C-delta-1 correlates with onset of neurogenesis. The C-delta-1 | | |
| CC | amino acid sequence was deduced from a cDNA clone (T58898) obt. | | |
| CC | from chick stage 4-6 embryos. A shorter version (W58877) of | | |
| CC | C-delta-1, lacking the 12 C-terminal amino acids of the longer | | |
| CC | version, was also isolated, and mouse (W11720) and human (W11721- | | |
| CC | 38) Delta-1 polypeptides have been identified. Delta-1 proteins | | |
| CC | can be used to treat or prevent disorders characterised by | | |
| CC | increased Notch activity, such as cervical, breast, lung or colon | | |
| CC | cancer, melanoma or seminoma, and nervous system disorders or to | | |
| CC | promote tissue regeneration and repair. | | |

| Seq | Sequence | 740 AA; |
|-----------------------|--------------------------------------------------------------------|---------------------------------------------|
| Query Match | 31.0%; | Score 1910; DB 21; Length 740; |
| Best Local Similarity | 50.1%; | Pred. No. 9,95e-151; |
| Matches | 263; Conservative | 78; Mismatches 168; Indels 16; Gaps |
| Db | 20 qvdsqgsvfklkqelvnhkqglslsmnccrccggppgagagqdcctffvclkhkyqasvsp | 79 |
| Qy | 19 QVHSGSGSEFLRKYSRSDNHGRDNERCCSGSDGATGCLCASCATRRVCAKHHQATIDT | 78 |
| Db | 80 eppccygaalfvjlgaansfsvpdgagadpafsmplrifqfvtvptgtsllaealhtdsp | 135 |
| Qy | 79 TSGCCTGYVITPILIGENSVNLTDQRPKNKFTNPIDQPFSEFSPGTFSLVEAMH-DTN | 133 |
| Db | 140 ddlittemperllsrlatqrhlavgeewsqdlhsgrtdlkysyrfvcdelhyegscavfc | 195 |
| Qy | 138 NSGNARTNKLLIQRLLVQOVLVESEEMTKNSESQYTLSEDFEFTVTCDLNLYSGCAKFC | 197 |
| Db | 200 rprdrfrfhtcgtgergekvcmprwkygqctepiclpqsgdehgfcdtpgeekcrvgyqr | 255 |
| Qy | 198 RPRDSEFGHSTSEFGELIICLWGGDVCCHPKCAKQG-E-HGHCIDPNQVCVCLGMKA | 255 |
| Db | 260 ydecirypgcrlhgtcggpwwqncqceqyqglfcnnqdlcythchhpkncgatcttqggsy | 319 |
| Qy | 256 LCNBEVLERNCHHGCKNRPWICINCENSGGLXKQNDLNYCNHNRKPNKGRCFNTGBGLY | 315 |
| Db | 320 tcsctpygtyscscelelnecda--npckngscctd-le-ns-yscctppgyfgxnclsa | 374 |
| Qy | 316 TCKCAPGYSGDDCENEIRYSCADAVNPCONGGTICIDEPHTKTYGKCHCANGMGCCEKY | 375 |
| Db | 375 mtccagpcctng-r-c-td-npdy-gysctcprplysgfneckkldycasspcanagacvd | 422 |
| Qy | 376 LTCSKRPCHQGCIRNVRPGLASKGGYCCDEP IGVSGPNCDOLDQDNCSPNPCLNGSGC-Q | 433 |
| Db | 430 lgnyslccqcgafgtrchddnvdcaafpcvngqfcdqgvndysctcoppnyngkncstpv | 488 |
| Qy | 435 -PSG-KCLICPAGFSTRETETNIDDLGHCQCGENGTCDIMVQNYKQCVPRGHGHCSSKY | 492 |
| Db | 490 srcehnpchgatachcrsnaryvccacarygglnqcfllpeppgpp | 534 |
| Qy | 493 DCLLRPCANGGTCLNLNDYQCTCRAGFTGKDCSVDIDECSSGP | 537 |
| RESULT | 4 | |
| ID | W18353 standard; protein; 723 AA. | |
| AC | W18353; | |
| DT | 11-FEB-1998 (first entry) | |
| DE | Proliferation and differentiation suppression polypeptide. | |
| KW | Proliferation; differentiation; suppression; human; delta-1; | |
| KW | serrate-1; blood cell; neuron; leukaemia; malignant tumour; | |
| OS | Immunosuppression. | |
| OS | Homo sapiens. | |
| FH | Key | |
| FT | Peptide | Location/Qualifiers |
| FT | | 1..21 |
| FT | Protein | /label= Signal |
| FT | | 22..723 |
| FT | | /label= Differentiation_suppression_protein |
| PN | WO9719172-A1. | |
| PD | 29-MAY-1997. | |
| PF | 15-NOV-1996; J03356. | |
| PR | 30-NOV-1995; JP-311811. | |
| PR | 17-NOV-1995; JP-299611. | |
| PA | (ASAH) ASAH1 KASEI KOGYO KK. | |
| PI | Itoh A, Sakano S; | |
| DR | WPI: 97-298110/27. | |
| DR | N-PSDB: T70174. | |
| PS | Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress | |
| PT | Proliferation and differentiation of undifferentiated human blood | |
| PT | cells | |
| PS | Claim 15: Page 77-82; 114pp; Japanese. | |
| CC | The present sequence represents a polypeptide which suppresses | |
| CC | proliferation and differentiation of undifferentiated cells such | |
| CC | as neurons and blood cells. The polypeptide may be used for the | |

| | Best Local Similarity 49.1%: | Pred. No. 1,44e-144: | |
|-------------------------------------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------|-----|
| Matches 256; Conservative 80; Mismatches 164; Indels 21; Gaps 17; | | | |
| Db | 1 | sgvfeklqgefwnkkgllgnncrcrga-gppp-c--acrtffvcclkhyaavspcpc | 56 |
| Oy | 23 | SGSFELRLKRYENSHDGRNDEGRCCSGESDQATGKLGCSKFRVCLKHQATIDRTTSQC | 82 |
| Db | 57 | tygsavtrvjayvdsfslpddgg-adsatsnrlrpfqfwtptgtsllaeahcdspdda | 115 |
| Oy | 83 | TYGDVDTPLIGENSVNLIDDAQRFONKGTNTDIOEPFSFWGTSLSLVEAMH-DTNSGN | 141 |
| Db | 116 | temperllsrlatqrphlrvgeewsqdlhsagrdlkysrffcvdehyhggcgvfrcpr | 175 |
| Oy | 142 | ARTNKLILGRLLVQOVLEVSSDMKTNKESQSYTSLIEDFRYTCIDNLNYSGCARFCPR | 201 |
| Db | 176 | dafightgergerkvcnppwkykqrypteplclpgcdeqhfcdkpygeckcrvvgwryr | 235 |
| Oy | 202 | DSFGHSTCSFETGEIICLTGMQGDYCHIRPKAKGC-E-HGCHDKPNQCCVCLGMKALCNE | 259 |
| Db | 236 | clrypgclhlyrcqcpwgcngcgegygllfcngdlhyctbhykpkagcatctntgsgysc | 295 |
| Oy | 260 | CYLEPNCIHGICNKPFWCTICNEGSGLYCNDLNYCTINHRKCNKGCTCFNGBELTYCK | 319 |
| Db | 296 | rpytygacelgldgcdp--spckngsactd-le-n-s-ysectcpbgfygkfcelsamtca | 350 |
| Oy | 320 | APGTSGDCCENBEIYSCDADVAPQNGGTCIDEPHTKIKYKHCANSGMKCEKXVLTCS | 379 |
| Db | 351 | dqpcfnng-r-c-sd-spdg-gysactrcpvygfnicekkldycsspscnagacvdlgda | 405 |
| Oy | 380 | DKPCHQGICRNVRPGLSGKGOGYCEPFIYGSFNCIDQLDNCSPNPCINGSCQPSGK- | 438 |
| Db | 406 | ylrcrqagfsgrhddndvdcasspcnagfcrigvndfctcprpytygncasparvce | 465 |
| Oy | 439 | --CICPAFSSTKCTCTNIDDLGHQCEGAGCTIDMKNQYRCQVPGFHGHCSKVDLCL | 496 |
| Db | 466 | hapcngatcheryhryycecarvygpnqcfllpelpbp | 506 |
| Oy | 497 | IRPCANGSTCLNNDYQCTCRAGTGRKDCVDDIDECSSP | 537 |
| RESULT | 7 | | |
| ID | W11720 | standard; Protein; 722 AA. | |
| AC | W11720: | | |
| DT | 28-APR-1997 | (first entry) | |
| DE | M-delta-1 polypeptide. | | |
| KW | M-delta-1; cell proliferation; nervous system disorder; | | |
| KW | tissue regeneration; Notch; cervix cancer; breast cancer; | | |
| KW | lung cancer; colon cancer; melanoma; seminoma; | | |
| OS | neurogenesis; therapy. | | |
| CS | Mus sp. | | |
| PN | W09701571-A1. | | |
| PD | 16-JAN-1997. | | |
| PE | 28-JUN-1996; U11178. | | |
| PR | 28-JUN-1995; US-000589. | | |
| PA | (IMCR) IMPERIAL CANCER RES TECHNOLOGY. | | |
| RA | (UYTA) UNIV YALE. | | |
| PI | Aravanijs-Tsakonas S, Gray GE, Henrique D, Ish-Horowitz D; | | |
| P1 | Lewis J; | | |
| DR | WPI: 97-100159/09. | | |
| DR | N-PSDB: T58899. | | |
| PT | New vertebrate Delta protein, DNA and antibodies - for treating and | | |
| PT | preventing cancer, nervous system disorders and for tissue | | |
| PT | regeneration | | |
| PS | Claim 4: Fig 8; 135pp; English. | | |
| CC | M-delta-1 polypeptide (W11720) is the mouse homologue of Drosophila | | |
| CC | Delta, a protein that binds to Notch protein. It is expressed | | |
| CC | primarily in presomitic mesoderm, the central and peripheral | | |
| CC | nervous systems, and kidney. Chick (W11719) and human (W11721- | | |
| CC | 38) Delta-1 polypeptides have also been identified. Delta-1 | | |
| CC | proteins can be used to treat or prevent disorders characterised by | | |
| CC | increased Notch activity, such as cervical, breast, lung or colon | | |
| CC | cancer, melanoma or seminoma, as well as nervous system disorders, | | |
| CC | and to promote tissue regeneration and repair. | | |

| Query Match | 29.6% | Score 1826; | DB 213; | Length 722; |
|-------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------|------------|-------------|
| Best Local Similarity 48.0%; | | Pred. No. 1,94e-143; | | |
| Matches 254; Conservative 87; | | Mismatches 166; | Indels 22; | Gaps 17; |
| DB | 14 | allcgvswsagvfejkqlgfevfkkgkllgnrnccrgss-gpp--c--acrtffvclkhya | 68 | |
| QY | 15 | TVIVVHSSGSFEELRLKRFSDNRDHRDNRGRCSSDSGACIGKCLGSKTFRCLKHYA | 74 | |
| DB | 69 | swspspeptcytsavhpvljvdsfslpdaag1-dpaefnplrfgf7wpqfsl1lea1h | 137 | |
| QY | 75 | TDITTSQCTYGDVITPILIGENSVMLDQARQNRKGFNPIDPFPSFMPSTSLYEAMH | 134 | |
| DB | 128 | tdspddlatemperlislrtqrlhltvgewsqdlhssgtrdlrlysfryfcdelyhygec | 187 | |
| QY | 135 | -DTNNSGNARNTKLLIOELVQOVLEVSSEWTKRKSQYTSLEYDFRVTCDLNYGSGC | 193 | |
| DB | 188 | svfcpridafghftcqdrgekcmdbpwkqgycctdplclpgcdqghycdkpgcekrvg | 247 | |
| QY | 194 | AKFCPRPDSGHSCTSGEIIICLTGMSDYCHIRPCAKGE--RHCHCKPNCQVCLG | 251 | |
| DB | 248 | wgryrdeccltyppgclhgtcqpwgqncqegwag1fcngdlnyctbkhpcrnatctng | 307 | |
| QY | 252 | WKGALCNCEVLEPNCIHGTCKNPMWTCICNBEWGGLYCNOIDLNTYCNHRPCKNGTCNTG | 311 | |
| DB | 308 | qgsyrcscrpytganclelvdec-ap-spcnkasctd-led-s-fscclpogfygkvc | 362 | |
| QY | 312 | ELLYLCKAPFSGSDCENEIYSCDADVPNCQNGCTCIDEPHRTGTYKCHCANGWSGKMC | 371 | |
| DB | 363 | elsamtcadpocfngg-r-c-sd-npdg-gytcchplqfsgfncckmldcgsspcnga | 417 | |
| QY | 372 | EKKVLTGCDKRCGICGNVNRPLGSKGQGYGCCPCPIGSSPNDDLDMCSPNCLNGG | 431 | |
| DB | 418 | kcvdlgnslrcqagfgyrcehdnvddcaaspcangctcrdsyndfscctppgytgkc | 477 | |
| QY | 432 | SCQPSGK---CICPFGSGTRCENINDDCLGHOCENGGTCDIMVNOYRCQCVPGFHGTHC | 488 | |
| DB | 478 | sapvrccehaphngatchtqgqymcecaagyygpnccqlpsepppg | 526 | |
| QY | 489 | SSKVDLCIRPCANGTCLNINDYQCTCRAGFGTKDCSVDIDECSSGP | 537 | |
| RESULT 8 | | | | |
| ID | R38305 | standard; Protein; 236 AA. | | |
| AC | R38305; | | | |
| DT | 30-NOV-1993 | (first entry) | | |
| DE | Sequence of a delta protein. | | | |
| KW | Delta; topotypic protein; family. | | | |
| OS | Drosophila melanogaster. | | | |
| PN | W09312141-A. | | | |
| PD | 24-JUN-1993. | | | |
| PR | 11-DEC-1991; 009240. | | | |
| PR | 11-DEC-1991; WO-009240. | | | |
| PA | (UYVA) UNIV YALE. | | | |
| PI | Artavanis-tsakonas S, Fleming RJ; | | | |
| DR | WPI; 93-214095/26. | | | |
| DR | N-PSDB; Q43911. | | | |
| PT | Purified serrate protein, nucleic acid and antibody used in | | | |
| PT | the study and manipulation of differentiation and other | | | |
| PT | physiological processes | | | |
| PS | Disclosure; Page 85; 119pp; English. | | | |
| CC | Delta encodes an approx. 100 kd protein (Delta denotes 'DIZM', the | | | |
| CC | protein product of the predominant zygotic and maternal | | | |
| CC | transcripts) that has nine EGF-like repeats within its | | | |
| CC | extracellular domain. Molecular studies have lead to the suggestion | | | |
| CC | that Notch and Delta constitute biochemically interacting elements | | | |
| CC | of a cell communication mechanism involved in early developmental | | | |
| CC | decisions. | | | |
| Sequence | 236 AA; | | | |
| Query Match | 28.9%; | Score 1784; | DB 7; | Length 236; |
| Best Local Similarity 100.0%; | | Pred. No. 8.50e-140; | | |

```
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mhwikclltaficftvivqvhsgsfelrlkyfsndhgrdnegrcsgesdgatgkclgs 60
Qy 1 mhwikclltaficftvivqvhsgsfelrlkyfsndhgrdnegrcsgesdgatgkclgs 60
Db 61 cktfrvclkhqyatidttscqtgydvitpiligensvmltdaqrnfqkftnpiqpfss 120
Qy 61 cktfrvclkhqyatidttscqtgydvitpiligensvmltdaqrnfqkftnpiqpfss 120
Db 121 svpgtfsllveahdtnnsqartnkllqirllvqvqlvsswsknkksesqytsleydf 180
Qy 121 svpgtfsllveahdtnnsqartnkllqirllvqvqlvsswsknkksesqytsleydf 180
Db 181 rvtcdlnyygscakrcprddsfghstcsetgeiicltgwqdychpckacgce 236
Qy 181 rvtcdlnyygscakrcprddsfghstcsetgeiicltgwqdychpckacgce 236

RESULT 9
ID W05835 standard; Protein; 1193 AA.
AC W05835;
DT 28-JAN-1997 (first entry)
DE Chick Serrate.
KW C-Serrate; Notch; cell differentiation; cell fate; tissue repair;
KW central nervous system; cancer; therapy; diagnosis.
OS Gallus sp.
FH Key
FT Location/Qualifiers
FT 1..1041
FT /label= Extracellular_domain
FT 1..5
FT /label= Sig_peptide
FT /note= "lacks the N-terminal portion owing to
FT truncation of the encoding cDNA clone"
FT 158..203
FT /label= DSL
FT /note= "region of homology with Drosophila Delta
FT and Serrate, predicted to mediate binding
FT with Notch"
FT 208..837
FT /label= ELR
FT /note= "epidermal growth factor-like repeat domain"
FT 208..238
FT /label= ELR1
FT 239..274
FT /label= ELR2
FT 275..313
FT /label= ELR3
FT 314..351
FT /label= ELR4
FT 352..390
FT /label= ELR5
FT 391..427
FT /label= ELR6
FT 428..464
FT /label= ELR7
FT 465..502
FT /label= ELR8
FT 503..540
FT /label= ELR9
FT 541..606
FT /label= ELR10
FT 607..644
FT /label= ELR11
FT 655..682
FT /label= ELR12
FT 683..721
FT /label= ELR13
FT 722..759
FT /label= ELR14
FT 760..797
FT /label= ELR15
FT 798..837
```

```
FT region /label= ELR16
FT 854..911
FT /label= Cysteine-rich_region
FT domain 1042..1066
FT /label= Transmembrane_domain
FT domain 1067..1193
FT /label= Intracellular_domain
FN W09627610-AL.
PD 12-SEP-1996.
PR 07-MAR-1995; U031172.
PR 07-MAR-1995; US-400159.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYA ) UNIV YALE.
PI Artavanis-Isakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
PI Lewis JH, Mann RS, Myat AM;
DR WPI; 96-425379/42.
DR N-PSDB; T40092.
PT Vertebrate Serrate protein and related DNA - used to treat or
PT prevent malignancies characterised by increased Notch activity.
PS Disclosure; Page 112-115; 161pp; English.
CC Chicken Serrate (W05835), or C-Serrate, is a ligand for the zygotic
CC neurogenic locus Notch and is believed to play a major role in
CC determining cell fates in the central nervous system. Its amino
CC acid sequence was deduced from a cDNA clone (T40092) obt'd. from an
CC optic explant cDNA library. C-Serrate is expressed in the central
CC nervous system, cranial placodes, nephric mesoderm, vascular
CC system, and limb bud mesenchyme.
SQ Sequence 1193 AA;

Query Match 27.9%; Score 1718; DB 19; Length 1193;
Best Local Similarity 43.1%; Pred. No. 4,46e-134;
Matches 239; Conservative 103; Mismatches 184; Indels 28; Gaps 19;

Db 1 qvasasgqfeilelsvqnvngvliqngnccdgtrnpdtkctrdcdctyfkvckeygsrv 60
Qy 19 QVHS-SGSFELRLKYFSNDHGRDNREGRCSSGESDGTGKCL-GSKTRFRVCLKHQYATI 76
Db 61 tagpccsfsgkstpvlgntfnlkys-r--nne-knrivipffawprsvllveaw-dy 115
Qy 77 DTTSQCTYGVITPILIGENSVNLTDAQRNFQKFTNPIQPFSSFGTSLIVEAWHDT 136
Db 116 ndhs-tnpdri-lekashsgminpsqwtlkhntgaahfeyqirvtcaehyqfgcnkf 173
Qy 137 NNSGNARTNKLILQRLVQVQLVSSSEWKNKSESYTSLEYDFRVTCDLNYGSGCAKF 196
Db 174 crprddffthhtcdqngktclegtwgpccnkaicrgcspkhgscstvpgecrqyqwg 233
Qy 197 CRPRDDSFHSTCSETGEIICLTGWQDYCHIPKACGCE--HGCHKPNQVCQLGWKG 254
Db 234 qycdkciphpqgvhgtcielpwqclcetnwggqlcdkldnycgthppclnggtsentgpd 293
Qy 255 ALNCEVLEPNCINGTCNKFTWTCNEGWLGYCNODLNTCTNHRPCKNGTCFNTGEGL 314
Db 294 yqcscpegysgqncelaeahclsd--pchngscl-e--tstqfecvcagwagptctdn 348
Qy 315 YTCAPGYSDDDCENEIYSCDADVNPCNGGTCIDDPHTKTKYKHCANGSKWCEEK 374
Db 349 iddcspnpcghgttcqd----l-v--dgfkicppqwtgktcqldanecgkpcvnansc 401
Qy 375 VLTCSDKPC-HQICRNVRPLGSKGOGYQCECPIGYSGPNCDLQDNCSNPNCINGSC 433
Db 402 rnlgsyydcitgwsghncdinordg-qcnggscrdlnvgyrcicspgyagdhcek 460
Qy 434 QP-SGK--CICPAGFSGTRCETNIDDLGHQCENGGTGTCIDMVNQYRCQCVPGFHTCSS 490
Db 461 dinecasnpcmngghcqddeingfclopagfsgnlcaldidycepnpcqagaqcfnlamd 520
Qy 491 KVDLCILIRPCANGGTCLNLANDYQCTCRAGFTGKDCSVDIDECSSGCHNGGTOMNRVNS 550
Db 521 yfncncpdyegknc 534
Qy 551 FECVCANGFRGKC 564
```

| ID | RESULT | 10 |
|----|--------------------------------------------------------------------|-----------------------------------------------------|
| AC | W05833; standard: Protein; 1218 AA. | |
| DT | 28-JAN-1997 (first entry) | |
| DE | Human Serrate-1 (HJ1). | |
| KV | Serrate-1; human jagged-1; HJ1; Notch; cell differentiation; | |
| KW | cell fate; central nervous system; cancer; tissue repair; therapy; | |
| KW | diagnosis; antibody. | |
| OS | Homo sapiens. | |
| EH | Key | |
| FT | domain | Location/Qualifiers |
| FT | peptide | 1..1067 |
| FT | domain | /label= Extracellular_domain |
| FT | peptide | 14..29 |
| FT | domain | /label= Sig_Peptide |
| FT | peptide | 185..229 |
| FT | peptide | /label= DSL |
| FT | peptide | /note= "region of homology with Drosophila Delta |
| FT | peptide | and Serrate, predicted to mediate binding |
| FT | peptide | with Notch" |
| FT | peptide | 234..896 |
| FT | peptide | /label= ELR |
| FT | peptide | /note= "epidermal growth factor-like repeat domain" |
| FT | peptide | 234..264 |
| FT | peptide | /label= ELR1 |
| FT | peptide | 265..299 |
| FT | peptide | /label= ELR2 |
| FT | peptide | 300..339 |
| FT | peptide | /label= ELR3 |
| FT | peptide | 340..377 |
| FT | peptide | /label= ELR4 |
| FT | peptide | 378..415 |
| FT | peptide | /label= ELR5 |
| FT | peptide | 416..453 |
| FT | peptide | /label= ELR6 |
| FT | peptide | 454..490 |
| FT | peptide | /label= ELR7 |
| FT | peptide | 491..528 |
| FT | peptide | /label= ELR8 |
| FT | peptide | 529..566 |
| FT | peptide | /label= ELR9 |
| FT | peptide | 567..598 |
| FT | peptide | /label= Partial_ELR |
| FT | peptide | 599..632 |
| FT | peptide | /label= Partial_ELR |
| FT | peptide | 633..670 |
| FT | peptide | /label= ELR10 |
| FT | peptide | 671..708 |
| FT | peptide | /label= ELR11 |
| FT | peptide | 709..747 |
| FT | peptide | /label= ELR12 |
| FT | peptide | 748..785 |
| FT | peptide | /label= ELR13 |
| FT | peptide | 786..823 |
| FT | peptide | /label= ELR14 |
| FT | peptide | 824..862 |
| FT | peptide | /label= ELR15 |
| FT | peptide | 863..879 |
| FT | peptide | /label= Partial_ELR |
| FT | peptide | 880..896 |
| FT | peptide | /label= Partial_ELR |
| FT | peptide | 1068..1089 |
| FT | peptide | /label= Transmembrane_domain |
| FT | peptide | 1090..1218 |
| FT | peptide | /label= Intracellular_domain |
| PN | W09627610-A1. | |
| PD | 12-SEP-1996. | |
| PF | 07-MAR-1996; U031172. | |
| PR | 07-MAR-1995; US-400159. | |
| PA | (IMCR) IMPERIAL CANCER RES TECHNOLOGY. | |
| PA | (UYTA) UNIV YALE. | |
| PA | Artavanis-Tsakonas S, Gray GE, Henrique DWP, Ish-Horowitz D; | |
| PI | Lewis JH, Mann RS, Myat AM; | |

[illegible]

```

FT Protein /label- Signal
FT 32..1218
FT /label- Differentiation_suppression_protein
PN W09719172-A1.
PD 29-MAY-1997. J03356.
PF 15-NOV-1996; JP-311811.
PR 30-NOV-1995; JP-299611.
PR 17-NOV-1995; JP-299611.
PA (ASAH ) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI: 97-298110/27.
DR N-PSDB; T70175.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 15; Page 83-91; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1218 AA;

Query Match 27.3%; Score 1680; DB 25; Length 1218;
Best Local Similarity 41.5%; Pred. No. 8,74e-131;
Matches 235; Conservative 108; Mismatches 194; Indels 29; Gaps 19;

Db 16 lllallc-alravcagsgqfeillsmqnvngelqncgccgarnpgdrktrdecdyt 74
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 7 LLTAFICIVIVQVH-SSGSFELRLKYFSDHGRDNEGRCCSGESDAGTKCL-GSKTR 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 75 fkvclkyqsrvtgagpcsfsgstpvigntfnlkas-rgnr---nrivlpsfawpr 130
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 65 FRVCLKHYQATIDTTSQCTYGVITPILGENSVNLTAQRFQNKGTNPQFPFSFSG 124
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 131 sytliveaw-dsmd-tvpdsi-iekashgminpsrqwtlikqntgvahfeyqirvtc 187
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 125 TFSLVEAWHDFTNSGNATNKLRLQVLLVQVLEVSSEWTKNSQYTSLEYDFRVT 184
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 188 ddyvvgfncfcrprddffghyacdqnqktcmegwmpcncraicrggcsphgskcl 247
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 195 DLYYSGCAKCRPRDSDFGHSTCSETGEIICLTGWQGYCHIPKCAKGE--HGCDK 242
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 248 pgdrcrgywgqglycdkclphpgcvhgicnepwqclcetnwggqlcdkldnycgthqpc 307
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 243 PNQVCQLGWKALGALNECVLEPNCIHGTCNKPTWCICNEGWGLYCNQDLYNCTNHRPK 302
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 308 nggtcsntgpkycscpegysgpnceiaehaclsd--pchnrgsc-ke--tslgfec 362
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 303 NGGTCFNTGEGLYTKCAPYSGDDCENIYSCDADVNPQNGGTCIDEPTHTKGYKCH 362
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 363 spgwtgptctnidcsppncshggtcqd---l-vng--fkvcppqwtgkctclidanc 415
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 363 ANWSGKWEKVEKVLKCDRPEC-HQGLCRNVRPLGSKGQGYQCECIGYSGPNCDQLDN 421
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 416 ceakpcvnaksknliasydcldpwmqgncdinindclg-qcndascdrlvngyrci 474
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 422 CSPNPNINGSCQP-SGK--CICPAGFSGTRCTNIDDCLGHCENGCGTCIDMVNYRCQ 478
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 475 cpbgvagdhceerdidecanspncgghcneinrfclptgfsnqlcldidycenpc 534
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 479 CVPFGHGTCHSSKVDLCILRCPANGGTCILNNDYQCTCRAGFTGDKCSVDIDECSSGPC 538
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 535 qnagcynrasdyfckcpdyegknc 560
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 539 HNGGTCMNRVNSFECVANGFRGKQC 564
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

RESULT 12
ID W18351 standard; protein: 1036 AA.
AC W18351;
DT 11-FEB-1998 (first entry)

```

DE Proliferation and differentiation suppression polypeptide.
KW Proliferation; differentiation; suppression; human; delta-1;
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
PN W09719172-A1.
PD 29-MAY-1997. J03356.
PF 15-NOV-1996; JP-311811.
PR 30-NOV-1995; JP-299611.
PR 17-NOV-1995; JP-299611.
PA (ASAH ) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI: 97-298110/27.
DR N-PSDB; T70175.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 5; Page 66-71; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1036 AA;

Query Match 27.2%; Score 1676; DB 25; Length 1036;
Best Local Similarity 42.1%; Pred. No. 1.94e-130;
Matches 231; Conservative 104; Mismatches 187; Indels 27; Gaps 17;

Db 1 sgqfeilellsmqnvngelqncgccgarnpgdrktrdecdyfkclykqysrvtaggp 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 23 SGSEFELRLKYFSDHGRDNEGRCCSGESDAGTKCL-GSKTRFVRLKHQYQATIDTSQ 81
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 61 csfgsgstpvigntfnlkas-rgnr---nrivlpsfawprsytlvew-dsmd-t 114
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 82 CTYGDVITPILGENSVNLTAQRFQNKGTNPQFPFSFSGPTFSLIVEAWHDFTNSGN 141
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 115 vqpdsi-iekashgminpsrqwtlikqntgvahfeyqirvtcdydygfcnkfcrpd 173
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 142 ARTNKLILRLQLVQVLEVSSEWTKNSQYTSLEYDFRVTCDLYYSGCAKCRPRD 201
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 174 dffghyacdqnqktcmegwmpcncraicrggcsphgskclpgdrcrgywgqglycdk 233
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 202 DSGHSTCSETGEIICLTGWQGYCHIPKCAKGE--HGCDKPNQVCQLGWKALGALNE 259
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 234 ciphpgcvhgicnepwqclcetnwggqlcdkldnycgthqpcinggtcsntgpkycsc 293
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 260 CVLEPNCIHGTCNKPTWCICNEGWGLYCNQDLYNCTNHRPKNGGTCFTNGSLYTCR 319
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 294 pegysgpnceiaehaclsd--pchnrgsc-ke--tslgfecscpgwtgptctnidcs 348
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 320 APGYSGDDCENIYSCDADVNPQNGGTCIDEPTHTKGYKCHANGWSGKMEKVEKVL 379
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 349 pnnchsggtcqd---l-vng--fkvcppqwtgkctclidaneceakpcvnaksknli 401
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 380 DKFC-HQGLCRNVRPLGSKGQGYQCECIGYSGPNCDQLDNCSNPNINGSCQP-SG 437
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 402 syvdcldpwmqgncdinindclg-qcndascdrlvngyrcicpbgvagdhceerdidec 460
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 438 K--CICPAGFSGTRCTNIDDCLGHCENGCGTCIDMVNYRCQVCPGFHGTCHSSKVDLC 495
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 461 aspnclngghcneinrfclptgfsnqlcldidycenpcnqagcynrasdyfckc 520
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 496 LIRPCANGGTCILNNDYQCTCRAGFTGDKCSVDIDECSSGPCMNRVNSFECVC 555
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 521 pedvegknc 529
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 556 ANGRGKQC 564
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

RESULT 13
ID W18352 standard; protein: 1187 AA.
AC W18352;

| ID | Accession | Score | DB | Length | Matches | Conservative | Mismatches | Indels | Gaps |
|-----------------------|----------------------------------------------------------------------|----------------------|----------------------|--------------|---------|--------------|------------|--------|------|
| DE | 11-FEB-1998 (first entry) | 27.2% | 1676; | 1187; | | | | | |
| DT | Proliferation; differentiation; suppression; human; delta-1; | 42.1% | Pred. No. 1,946-130; | | | | | | |
| KW | serrate-1; blood cell; neuron; leukaemia; malignant tumour; | 104; | Mismatches 187; | | | | | | |
| KW | Immunosuppression. | | | | | | | | |
| OS | Homo sapiens. | | | | | | | | |
| PN | W09719172-A1. | | | | | | | | |
| PD | 29-MAY-1997. | | | | | | | | |
| PF | 15-NOV-1996; J03356. | | | | | | | | |
| PR | 30-NOV-1995; JP-311811. | | | | | | | | |
| PR | 17-NOV-1995; JP-299611. | | | | | | | | |
| PA | (ASAH) ASAH KASEI KOGYO KK. | | | | | | | | |
| PI | Itch A. Sakano S; | | | | | | | | |
| DR | WPI: 97-298110/27. | | | | | | | | |
| PT | Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress | | | | | | | | |
| PT | proliferation and differentiation of undifferentiated human blood | | | | | | | | |
| PT | cells | | | | | | | | |
| PS | Claim 6; Page 71-76; 114pp; Japanese. | | | | | | | | |
| CC | The present sequence represents a polypeptide which suppresses | | | | | | | | |
| CC | proliferation and differentiation of undifferentiated cells such | | | | | | | | |
| CC | as neurons and blood cells. The polypeptide may be used for the | | | | | | | | |
| CC | prevention and control of disorders involving undifferentiated | | | | | | | | |
| CC | cells, such as leukemia and malignant tumours, and improvement of | | | | | | | | |
| CC | blood formation, e.g. after immunosuppression. | | | | | | | | |
| SQ | Sequence 1187 AA; | | | | | | | | |
| Query Match | 27.2%; | Score 1676; | DB 25; | Length 1187; | | | | | |
| Best Local Similarity | 42.1%; | Pred. No. 1,946-130; | | | | | | | |
| Matches 231; | Conservative 104; | Mismatches 187; | Indels 27; | Gaps 17; | | | | | |
| Db | 1 sgqfcllmsngvngelqngnccggarppdrctdeclyfkvcokvsgvtagp 60 | | | | | | | | |
| Oy | 23 SGSEFLRKRTYSNDGRDNEGRCGSESDGATGKL-GSCKTRRVCKKHQATIDTTSQ 81 | | | | | | | | |
| Db | 61 csfgsgspivlgunfnlkaas-rgndr--nriylpfafazsytliveaa-dssnd-t 114 | | | | | | | | |
| Oy | 82 CTYGVVITPILIGENVNLTDNRQPNKGFTNPIDPFPSFSPGFFSLVEAMHDTNNSGN 141 | | | | | | | | |
| Db | 115 vqpdsi-lekashsgmlnpsrqwtlkqntvahnfeqydlvtcdyygfigcnkfcprd 173 | | | | | | | | |
| Oy | 142 ARTNKLILQRLVQVLEVSSEMKTKNSEOYTSLEDFRVTCDLNLYGSGCAKFRPD 201 | | | | | | | | |
| Db | 174 dffgryaaddqngntkcmegmmpreataicqgscspkhgseklpgrdcrcqywgylcdk 233 | | | | | | | | |
| Oy | 202 DSFGSTSEGEIICLIGWGDYCHIRKAKGE--HGHCDFKNQCVCQLGMKGALCNE 259 | | | | | | | | |
| Db | 234 ciphbgcngicnepwgclicetnuygqldkldlycghqpcnlngscsnlprpdkyqsc 293 | | | | | | | | |
| Oy | 260 CVLEBNCNCHGICNKPWCICINMGWGLCNDLWNTCHRRCKNKGICFNTGBELYTKCK 319 | | | | | | | | |
| Db | 294 pegysgpnccelaethacd--pchnrgsc-ke--tsjgfecescpwygtptcsnldcs 348 | | | | | | | | |
| Oy | 320 APGYSGDCDENIEYSCDADVNPCKGNGCTCIDEPHTKTKGCHCANGSGMKCEKRVLTCS 379 | | | | | | | | |
| Db | 349 pnnchgstlqcd---l-vng--fkvcpppqtgtlqgladneccahprvnaekschlia 401 | | | | | | | | |
| Oy | 380 DKPC-HGICIRKVRPGLSKGQGYQCECPIDYSSPNCDDLQDNCSPMPCINGSGCQCP-SG 437 | | | | | | | | |
| Db | 402 sydcclpgmwmgqndlnadclg-qcndascdrlvnygrcicpppyaqdchcrdiadc 460 | | | | | | | | |
| Oy | 438 K-CLCPAGFGSGTKETNIDDCLEHGCGNGSTCIDMWNYQYQCQVPEFHGHCSSKYDL 495 | | | | | | | | |
| Db | 461 asnpclngghcneinfqcldptfgisgnlqldldycenpncqngaqcyraasdyfck 520 | | | | | | | | |
| Oy | 466 LIRPCANGCTCLNINNDYQCTCRAGFTGKDCSDVIDEASSPCCHNGGFCMNRVNSFCVC 555 | | | | | | | | |
| Db | 521 pedyeqknc 529 | | | | | | | | |
| Oy | 556 ANGFGRKQC 564 | | | | | | | | |

| | | |
|----|--------------------------------------------------------------------|-----------------------------------------------------|
| AC | W05834;1997 | (first entry) |
| DE | Human Serrate-2 (HJ2). | |
| KW | Serrate-2; human jagged-2; HJ2; Notch; cell differentiation; | |
| KW | cell fate; central nervous system; cancer; tissue repair; therapy; | |
| KW | diagnosis; antibody. | |
| OS | Homo sapiens. | |
| PH | Key | Location/Qualifiers |
| FT | domain | 1..912 |
| FT | | /label= Extracellular_domain |
| FT | | /note= "a deletion in the encoding cDNA clone |
| FT | | results in loss of part of the Serrate-2 |
| FT | | signal peptide and beginning of the DSL |
| FT | domain | 26..70 |
| FT | | /label= DSL |
| FT | | /note= "region of homology with Drosophila Delta |
| FT | | and Serrate, predicted to mediate binding |
| FT | | with Notch" |
| FT | domain | 75..735 |
| FT | | /label= ELR |
| FT | region | /note= "epidermal growth factor-like repeat domain" |
| FT | | 75..105 |
| FT | | /label= ELR1 |
| FT | region | 106..140 |
| FT | | /label= ELR2 |
| FT | region | 141..180 |
| FT | | /label= ELR3 |
| FT | region | 181..218 |
| FT | | /label= ELR4 |
| FT | region | 219..256 |
| FT | | /label= ELR5 |
| FT | region | 257..294 |
| FT | | /label= ELR6 |
| FT | region | 295..331 |
| FT | | /label= ELR7 |
| FT | region | 332..369 |
| FT | | /label= ELR8 |
| FT | region | 370..407 |
| FT | | /label= ELR9 |
| FT | region | 408..435 |
| FT | | /label= Partial_ELR |
| FT | region | 436..469 |
| FT | | /label= Partial_ELR |
| FT | region | 470..507 |
| FT | | /label= ELR10 |
| FT | region | 508..545 |
| FT | | /label= ELR11 |
| FT | region | 546..584 |
| FT | | /label= ELR12 |
| FT | region | 585..622 |
| FT | | /label= ELR13 |
| FT | region | 623..660 |
| FT | | /label= ELR14 |
| FT | region | 664..701 |
| FT | | /label= ELR15 |
| FT | region | 702..718 |
| FT | | /label= Partial_ELR |
| FT | region | 719..735 |
| FT | | /label= Partial_ELR |
| FT | domain | 913..933 |
| FT | | /label= Transmembrane_domain |
| FT | domain | 934..1257 |
| FT | | /label= Intracellular_domain |
| PN | M09627610-A1. | |
| PD | 12-SEP-1996. | |
| PF | 07-MAR-1996; U03172. | |
| PR | 07-MAR-1995; US-400159. | |
| PA | (IMCR) IMPERIAL CANCER RES TECHNOLOGY. | |
| PA | (UYVA) UNIV YALE. | |
| PI | Aravanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D; | |
| PI | Lewis JH, Mann RS, Myat AM; | |
| DR | WPI; 96-425379/42. | |

DR N-PSDB; W05834.
 PT Vertebrate Serrate protein and related DNA - used to treat or
 PT prevent malignancies characterised by increased Notch activity.
 PS Claim 5; Page 104-107; 161pp; English.
 CC Human Serrate-1 (W05833) and human Serrate-2 (W05833) are ligands
 CC for the zygotic neurogenic locus Notch, and are believed to play a
 CC major role in determining cell fates (differentiation) in the
 CC central nervous system. Their amino acid sequences were deduced
 CC from cDNA clones (see also T40090-91) isolated from human foetal
 CC brain cDNA libraries. The proteins, antibodies raised to them,
 CC and encoding nucleic acids can be used in the detection of
 CC Serrate sequences and in the treatment of disorders of cell fate
 CC or differentiation, partic. cancer, nervous system disorders
 CC and in tissue repair or regeneration.
 SQ Sequence 1257 AA;

Query Match 21.8%; Score 1346; DB 19; Length 1257;
 Best Local Similarity 43.5%; Pred. No. 6.38e-102;
 Matches 177; Conservative 73; Mismatches 138; Indels 19; Gaps 11;

Db 8 wkslhsgvhahlelqirvrcdenyysatcnkforprndffghytcddqygnkacmdgmg 67
 QY 164 WTKNKSEQVTSLEYDFRVTCDLNYGSGCAKFCRPRDSDFGHSTCSETGEIICLTGWQ 223
 Db 68 keckeavckgcnllhggctvpccrscyqggrfdecypgpcvhgscvepqncet 127
 QY 224 DYCHIPKAKGCE--HGCDKPNQCVQGLGWKALCNECVLEPNCIHGTCKNKPWTCICNE 281
 Db 128 nwgllcdkdlncygshhpcntngtcinaeepdyrctcpdpgysgrncekaehac-ts-np 185
 QY 282 GWGGLYCNQDLNCTNHRPCKNGTCTGEGLYTCKCAPGYSDDCENIYSCDADVP 341
 Db 186 canggschevp----sgfechpcpswgptcaldidcasnpcaagtc--vdq-v-----d 235
 QY 342 CQNGGTCIDEPHTKTKYKCHANGSGKMEKVLTCSDKPCHQG-ICRNVRLGLSGKQ 400
 Db 236 gfecicpeqvqatclldaneceqgkclnafsctnllggyvcdipgwkinchinvc 295
 QY 401 GYOCEPIGYSGPCNDLQLDNCSFNPCINGSCQP--SGK-CICPAGFSGRCTETNIDDC 457
 Db 296 rg-qcghgtgckdlvngyqvcprfggrhcelrdkcaspsghgclcedladgfhc 354
 QY 458 LGHQCEGNGTCDIMVNOYRCQCPGPHGTHCSSKVDLCLIRPCANGTCLNLDNYQCTC 517
 Db 355 pggfslpcevdlcepsrcngarcynlegdyycapddfggknc 401
 QY 518 RAGFTGKDCSDVIDECSSGPCNNGTGMNRVNSFECVCANGFRGKQC 564

RESULT 15

ID R38304 standard; Protein; 1404 AA.
 AC R38304;
 DT 30-NOV-1993 (first entry)
 DE Sequence of a serrate protein.
 KW Serrate; topothymic protein; family.
 OS Drosophila melanogaster.
 PN W0931241-A.
 PD 24-JUN-1993.
 PF 11-DEC-1991; U09240.
 PR 11-DEC-1991; WO-U09240.
 PA (UYVA) UNIV YALE
 PI Artavanis-tsakonas S, Fleming RJ;
 DR N-PSDB; Q43910.
 DR Purified serrate protein, nucleic acid and antibodies - used in
 PT the study and manipulation of differentiation and other
 PT physiological processes
 PS Claim 4; Pages 74-80; 119pp; English.
 CC Two Drosophila genomic phage libraries were screened and recombinant
 CC clones were isolated. The cDNAs in lambda gt10 were isolated from an
 CC early pupal library. The C1 cDNA was isolated from an early pupal
 CC library. Subsequently the C3 cDNA was isolated using the 5', 700 bp
 CC terminal fragment of the C1 cDNA as probe. The complete 5561bp

CC sequence of DNA of the Drosophila Serrate protein was derived from
 CC C1 and C3 cDNAs (Q43910). The deduced protein product appears to be
 CC a transmembrane protein. AAs 51-80 represent the likely signal
 CC peptide; aas 542-564 represent potential membrane associated region;
 CC aas 1221-1245 represent the putative transmembrane domain.
 SQ Sequence 1404 AA;

Query Match 20.4%; Score 1258; DB 7; Length 1404;
 Best Local Similarity 34.8%; Pred. No. 2.42e-94;
 Matches 207; Conservative 143; Mismatches 199; Indels 46; Gaps 35;

Db 63 cnliali-lillvhkisaagfelleileisntshllngyccmpaelratktlqcspt 121
 QY 6 CLLTAFICFVIVQ-VHSSGSFELRLKXFNDRHGRDNEGRCCSGESDGTGKCLG-S-K 62
 Db 122 tafrlclkeyqtteqgasirrcsfngnatktllgssfvlsdp-----g-vgaivlpft 174
 QY 63 TRPRVCLKHQATID-TT-SQ-CTYGVITPILGENSEVNLTDQRFQNGFTNPDIQFPFS 119
 Db 175 frwtksflllea-lmynitsypdaerlieetsysgilp-spewktldhigrnaritry 232
 QY 120 FSWPGTFSLIVEAHHDTNNSGNARTKLLIQLRLVQVLEVSSEWTKNKSQYTSLEYD 179
 Db 233 vrvqavtynttcttferprddgfhycgseqgkiclngwqgvnceaeackagcdpvh 292
 QY 180 FRVTCDLNYGSGCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPKAKGCE--H 237
 Db 293 gkdrpgececrpgwrgplcnemvypgckhssngsawkcvcdtnwggllcdqdlncf 352
 QY 238 GHCDKPNQCVQGLGWKALCNECVLEPNCIHGTCTN-KPWTCICNEGWGLYCNQDLNVC 296
 Db 353 thepckhggtcentapdkyrtcaeglsgeqceivhepc-a-trpcrngtcttkttsnr 410
 QY 297 NHRPCKRGGTCTFNTGEGLYTCKCAPGYSDDCENIYSCDADVPNCPONGGTC-IDEPH-T 354
 Db 411 qaqvyrtshgrsnmrgprvrrssmrsldhrpeggalngsssglvsqslqlgqlapd 470
 QY 355 KTG-YK-CH-CAN-GWSGKWCCE-KVLTC--SDRPHQGICRNVRLGLS---KGQ-G-- 401
 Db 471 ftdcaagwtgptceinidecagpgcehggtctldliffreccpewhgdvcvqdvnece 530
 QY 402 YQCEPIGYSGPCNDLQLDNCSFNPCINGSGC-QP-SG-KCICPAGFSGRCTETNIDDC 458
 Db 531 aphsagiaaalltttataignslstallaaltsavastsiaigpcinakecrnpgs 590
 QY 459 GHQCEN-GGTCTI--DMVNOYR-CQ-CVPFGHGTCHCSSKVDLCL-IRPCANGTCLNLD 512
 Db 591 facickeqvgvtcaenldcvg-gccrngatcidlvndycacagsgftgrdcetd 644
 QY 513 YQCTCRAGFTGKDCSDVIDECSSGPCNNGTGMNRVNSFECVCANGFRGKQCDDE 567

Search completed: Fri Jun 12 11:50:28 1998

Job time : 43 secs.

CC TOPOLOGY: unknown
CC MOLECULE TYPE: Protein
SQ SEQUENCE 833 AA; 88612 MW; 3155492 CN;

Query Match 100.0%; Score 6164; DB 1; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWIKLLAFICFTVIVQVHSSGSEFLRLKYFSDHGRNDRCCSGESDGTGKGLGS 60
QY 1 MHWIKLLAFICFTVIVQVHSSGSEFLRLKYFSDHGRNDRCCSGESDGTGKGLGS 60

Db 61 CKTRFRVCLKHQATIDTTSQCTGYDVTTPILGNSVNLDAQRFQKNGFTNPQFFSF 120
QY 61 CKTRFRVCLKHQATIDTTSQCTGYDVTTPILGNSVNLDAQRFQKNGFTNPQFFSF 120

Db 121 SWPGTFSLIVAEWHDTNNSGNARTKLLIQRLVQVLEYSSEWKTNKSQYTSLEYDF 180
QY 121 SWPGTFSLIVAEWHDTNNSGNARTKLLIQRLVQVLEYSSEWKTNKSQYTSLEYDF 180

Db 181 RVTCDLNLYGSGCAKFCRPRDDSFHSTCTSETGEIICLTGWQDYCHIPKCAKCEHGHC 240
QY 181 RVTCDLNLYGSGCAKFCRPRDDSFHSTCTSETGEIICLTGWQDYCHIPKCAKCEHGHC 240

Db 241 DKPNQVCQGLGWGALNECVLEPNCIHGTCKNKPWTICNEGWGGLYCNQDLNCTNHRP 300
QY 241 DKPNQVCQGLGWGALNECVLEPNCIHGTCKNKPWTICNEGWGGLYCNQDLNCTNHRP 300

Db 301 CKNGGTCFNTGEGLYTKCAPYSGDDCENEIYSCDADVPNCQNGGTCIDEPHTKTYKC 360
QY 301 CKNGGTCFNTGEGLYTKCAPYSGDDCENEIYSCDADVPNCQNGGTCIDEPHTKTYKC 360

Db 361 HCANGSGKCEKVLTCSDPKPHQICRNVRPGLSGKGGYQCECPIGYSGPNCDQLD 420
QY 361 HCANGSGKCEKVLTCSDPKPHQICRNVRPGLSGKGGYQCECPIGYSGPNCDQLD 420

Db 421 NCSPNCPINGSCQPSKICPAGFSTCTETNTDCLGHQCENGGTCIDMVNRYRCQV 480
QY 421 NCSPNCPINGSCQPSKICPAGFSTCTETNTDCLGHQCENGGTCIDMVNRYRCQV 480

Db 481 PGFPGTHCCKSVLDLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVYDECSGGPCHN 540
QY 481 PGFPGTHCCKSVLDLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVYDECSGGPCHN 540

Db 541 GGTGCMNRVNFECVANGFCRGKQCDSEYSDVTPDAHOYGAATQARADGLTNAQVLI 600
QY 541 GGTGCMNRVNFECVANGFCRGKQCDSEYSDVTPDAHOYGAATQARADGLTNAQVLI 600

Db 601 FSVAMPLVAVIAACVFCMKRKRRAQEKDDAEARKQNEONAVATMHNGSGVGVALASA 660
QY 601 FSVAMPLVAVIAACVFCMKRKRRAQEKDDAEARKQNEONAVATMHNGSGVGVALASA 660

Db 661 SLGKGTGNSGLTFDGGNPNIKTNDKSVNNICASAAAAAADADECLMYGGYVASV 720
QY 661 SLGKGTGNSGLTFDGGNPNIKTNDKSVNNICASAAAAAADADECLMYGGYVASV 720

Db 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLHMRGSPAGSSAKGASGGPGAAEGRISVL 780
QY 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLHMRGSPAGSSAKGASGGPGAAEGRISVL 780

Db 781 GEGSYCSQRWPSLAAAGVAGACSSQLMAAASAGSGAGTAQOQSVVCGTPTHM 833
QY 781 GEGSYCSQRWPSLAAAGVAGACSSQLMAAASAGSGAGTAQOQSVVCGTPTHM 833

RESULT 2
ID US-08-264-534-3 STANDARD; PRT; 203 AA.
XX xxxxxx
AC xxxxxx
DT 01-JAN-1900
XX Sequence 3, Application US/08264534.

XX Sequence 3, Application US/08264534
CC Patent No. 5648464
CC GENERAL INFORMATION:
CC APPLICANT: Artavanis-Tsakonas, Spyridon et al.
CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
CC TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA: US/08/264,534
CC APPLICATION NUMBER: US/08/264,534
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/695,189
CC FILING DATE: 03-MAY-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 7326-004
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 8698864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 203 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC SEQUENCE 203 AA; 22840 MW; 218129 CN;

Query Match 24.6%; Score 1514; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.06e-113;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GSFELRLKYFSDHGRNDRCCSGESDGTGKGLGCKTRFRVCLKHQATIDTTSOCT 60
QY 24 GSFELRLKYFSDHGRNDRCCSGESDGTGKGLGCKTRFRVCLKHQATIDTTSOCT 83

Db 61 YGDVITPILGNSVNLDAQRFQKNGFTNPQFFSFQSWPGTFSLIVAEWHDTNNSGNAR 120
QY 84 YGDVITPILGNSVNLDAQRFQKNGFTNPQFFSFQSWPGTFSLIVAEWHDTNNSGNAR 143

Db 121 TNKLLIQRLVQVLEYSSEWKTNKSQYTSLEYDFRVTCDLNLYGSGCAKFCRPRD 180
QY 144 TNKLLIQRLVQVLEYSSEWKTNKSQYTSLEYDFRVTCDLNLYGSGCAKFCRPRD 203

Db 181 FGHSTCSETGEIICLTGWQDYC 203
QY 204 FGHSTCSETGEIICLTGWQDYC 226

RESULT 3
ID US-08-457-135-2 STANDARD; PRT; 383 AA.
XX xxxxxx
AC xxxxxx
DT 01-JAN-1900
XX Sequence 2, Application US/08457135.

GENERAL INFORMATION:
APPLICANT: LABORDA, Jorge
TITLE OF INVENTION: Delta-Like Gene Expressed In
Neuroendocrine Tumors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/597,545
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/989,537
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
STRAIN: Mouse Dlk
SEQUENCE 385 AA; 41320 MW; 772804 CN;

Query Match 10.6%; Score 654; DB 1; Length 385;
Best Local Similarity 37.0%; Pred. No. 3.63e-42;
Matches 85; Conservative 50; Mismatches 82; Indels 13; Gaps 10;
Db 26 CDPPCDPQYGFCEADNVCRCHVWGEGPLCDKCVTAPGCVNGVCKEPQWQICKDGDWQKFC 85
QY 231 CAKGCE-H-GHCDKPNQVCQLGKWCALCNCEVLEPNCIHGTCNKPWTCICNEGWGGLYC 288
Db 86 EIDVRACTS-TPCANNGTCVDLEKGOYECSTPGFSGKDCQHKAGPCVINGSPCQHGAC 144
QY 289 NQDLNYCTNHRPCKNGGTCFNTGEGLYTCKCAPGYSGDDCENEIYSCDADVNPQNGGTC 348
Db 145 VDDEGOASHASCLCPGFGNFC--IVA-ATNSCTPNPCEN--DGVCTDIGDFRCRCP 199
QY 349 IDEPHTKTYKCHCANGSGKMCCEKVLTCSDKPCHQGICRNVRPGL-GSKGGGYQCECP 407
Db 200 AGFVDTKTSRPVNSCASPQNGGTCLQHTQVSEFCLCKPPFMGPTCAKK 249
QY 408 IGYSGPNCDLQDNCSPNCGGSC-QPS--G-KCICPAGFSGTRCETN 453

RESULT 6
ID PCT-US-08-457-135-1 STANDARD; PRT; 385 AA.
XX
AC
XX
XX
DT 01-JAN-1900
XX
DE Sequence 1, Application US/08457135.
XX
CC Sequence 1, Application US/08457135
CC Patent No. 5644031
CC GENERAL INFORMATION:

APPLICANT: LABORDA, Jorge
TITLE OF INVENTION: Delta-Like Gene Expressed In
Neuroendocrine Tumors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,135
FILING DATE: 01-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/989,537
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/304/NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
STRAIN: Mouse Dlk
SEQUENCE 385 AA; 41320 MW; 772804 CN;
Query Match 10.6%; Score 654; DB 1; Length 385;
Best Local Similarity 37.0%; Pred. No. 3.63e-42;
Matches 85; Conservative 50; Mismatches 82; Indels 13; Gaps 10;
Db 26 CDPPCDPQYGFCEADNVCRCHVWGEGPLCDKCVTAPGCVNGVCKEPQWQICKDGDWQKFC 85
QY 231 CAKGCE-H-GHCDKPNQVCQLGKWCALCNCEVLEPNCIHGTCNKPWTCICNEGWGGLYC 288
Db 86 EIDVRACTS-TPCANNGTCVDLEKGOYECSTPGFSGKDCQHKAGPCVINGSPCQHGAC 144
QY 289 NQDLNYCTNHRPCKNGGTCFNTGEGLYTCKCAPGYSGDDCENEIYSCDADVNPQNGGTC 348
Db 145 VDDEGOASHASCLCPGFGNFC--IVA-ATNSCTPNPCEN--DGVCTDIGDFRCRCP 199
QY 349 IDEPHTKTYKCHCANGSGKMCCEKVLTCSDKPCHQGICRNVRPGL-GSKGGGYQCECP 407
Db 200 AGFVDTKTSRPVNSCASPQNGGTCLQHTQVSEFCLCKPPFMGPTCAKK 249
QY 408 IGYSGPNCDLQDNCSPNCGGSC-QPS--G-KCICPAGFSGTRCETN 453

RESULT 7
ID PCT-US91-09055-2 STANDARD; PRT; 1480 AA.
XX
AC
XX
XX
DT 01-JAN-1900
XX
DE Sequence 2, Application PC/TUS9109055.
XX
CC Sequence 2, Application PC/TUS9109055
CC GENERAL INFORMATION:
CC APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
CC TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
CC NUMBER OF SEQUENCES: 9

Db 178 CERPMPKNCSEAPAC-PEN-CLNGLCYRA-KCICEBGTGEDCSQARCPSPCNDQKCV 234
Cc 184 CDLNYTSSGCAK-FCRPHDSFGHSTCSETEIICLTGMQGGCHIPKACGC-BHGCD 241
Cc 235 D-GVCFEFGYTGDCGE-ELCPHCGFHRCVGR-CVCHGFTGEDCNEPL--CPNN- 288
Cc 242 KPNOCVCLGKAGALCNECVLEPN-C-THGICNKPWTICLNGMGLYCNDLNTCTNHR 299
Db 289 -CHNRGRVDN-E---CVDEGYTGEDCGELICPNDFDRGRCLNG-TCECEBGTGED 341
Cc 300 PCKNGGCFNTEGELYTKCAPGSGDDCENELYSCLA-DVNPQNGGCI-DEHFTKG 357
Cc 342 CGELTCNNKNGRCENGCLVCHEGFTVGDSCQARCPKTCNNRCVDC-RCVCHGYL 400
Cc 358 Y-KCHCANGMSGK-MCEKRYLTCSL---KPCHOICRNVRPGLSGKGGYCECPIGIS 411
Db 401 GEDCG-EL-RC-PNDCHNRGCI-NGOCVDEGTGEDCG-ELR-C-PNDQQRGCI- 452
Cc 412 GPNCDLQDNCSPPNCINGSCQPSGKCTCPAGFSGTRETINIDCLHQCENGGTCLDM 471
Cc 453 -GQ--CECHGEFTGEDCGE-LR-CP-NDNCHGRVCN--G-QCVDEGYTGEDCG-EL- 500
Cc 472 VNQRCQCFPHGTHGSKYDCLIRPCANGTCLNNDYQCTCRAGFTGKDCSVID 531
Db 501 RCPND-CHNRGR--VEG-RCVCDNGFMGEDCGE 530
Cc 532 ECSSGPGHNGTGMNRVNSFEVCANGFRGKQCD 566
Cc 532 ECSSGPGHNGTGMNRVNSFEVCANGFRGKQCD 566

RESULT 11
ID US-08-264-534-4 STANDARD: PRT: 199 AA.

AC xxxxxx

DT 01-JAN-1900

DE Sequence 4, Application US/08264534.

XX Sequence 4, Application US/08264534

CC Patent No. 5648464

CC GENERAL INFORMATION:

CC APPLICANT: Artavanis-Tsakonas, Spyridon et al.

CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains

CC TITLE OF INVENTION: In Topolythmic Proteins, And Methods Based Thereon

CC NUMBER OF SEQUENCES: 34

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pennie & Edmonds

CC STREET: 1155 Avenue of the Americas

CC CITY: New York

CC STATE: New York

CC COUNTRY: U.S.A.

CC ZIP: 10036

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/264,534

CC FILING DATE:

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/695,189

CC FILING DATE: 03-MAY-1991

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Mistock, S. Leslie

CC REGISTRATION NUMBER: 18,872

CC REFERENCE/DOCKET NUMBER: 7326-004

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 212 790-9090

CC TELEFAX: 212 869864/9741

CC TELEX: 66141 PENNIE

CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 199 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC SEQUENCE 199 AA; 21952 MW; 208406 CN;

Query Match 7.3%; Score 452; DB 1; Length 199;
Best Local Similarity 37.0%; Pred. No. 7.36e-26;
Matches 77; Conservative 51; Mismatches 66; Indels 14; Gaps 9;
Db 1 GNFELILEISTNSHLNGYCCGPAELRAKTKGSCCTTAFLRLKEYOTTEGAS 60
Cc 24 GSFELRLKXFSNDHGRDNRCCSGSDGATKCLG-S-CKTRFVCLKHQATID-TT- 79
Cc 61 STGCSFPGNATTKIIGSSPVLSDP-----G-VGATVLPETFRMTRKSFLLIOA-LDMYN 112
Cc 80 SO-CYGDVITPLILENSVNLTDQAFONKGTNPQFPFSFSPGTFSLIVEAMHDINN 138
Cc 113 TSYPAERLIEFTSYSGVILP-SPEKTLDHIGRNARITYRVVOCAYVYNTCTTFCR 171
Cc 139 SGNATNKLILQRLVQVYLEVSSSEKTKKSSQYTSLEYDFRYVCDLNYGSGCAKPCR 198
Db 172 PRDQFGHAGCSBQKCLNGMGVNC 199
Cc 199 PRDQFGHAGCSBQKCLNGMGVNC 199
Cc 199 PRDQFGHAGCSBQKCLNGMGVNC 199

RESULT 12
ID PCT-US95-02251-3 STANDARD: PRT: 1251 AA.

AC xxxxxx

DT 01-JAN-1900

DE Sequence 3, Application PC/TUS9502251.

XX Sequence 3, Application PC/TUS9502251

CC GENERAL INFORMATION:

CC APPLICANT:

CC TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE

CC TITLE OF INVENTION: CELLS

CC NUMBER OF SEQUENCES: 18

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Arnold, White & Durkee

CC STREET: P.O. Box 4433

CC CITY: Houston

CC STATE: Texas

CC COUNTRY: United States of America

CC ZIP: 77210

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: Patentin Release #1.0, Version

CC SOFTWARE: #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/02251

CC FILING DATE:

CC CLASSIFICATION: CONCURRENTLY HERewith

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/316,650

CC FILING DATE: 30-SEP-1994

CC CLASSIFICATION:

CC APPLICATION NUMBER: US 08/199,780

CC FILING DATE: 18-FEB-1994

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Parker, David L.

CC REGISTRATION NUMBER: 32,165

CC REFERENCE/DOCKET NUMBER: UMIC009P--


```
QY 301 CKNGTCFNTGEGLYCKCAPGYSGDDCENEIYSCDADVNPONGGTCIDEPHTKTYKC 360
Db 361 HCANGSGMKCEKVLTCSDKPCCHQICRNVRPGLSGKGGYQCECPICYSGPNCDLQLD 420
QY 361 HCANGSGMKCEKVLTCSDKPCCHQICRNVRPGLSGKGGYQCECPICYSGPNCDLQLD 420
Db 421 NCSPNPCIINGGSCQPSGKICIPAGSGTRCETNIDCLGHQCENGGTCIDMNVQYRCQCV 480
QY 421 NCSPNPCIINGGSCQPSGKICIPAGSGTRCETNIDCLGHQCENGGTCIDMNVQYRCQCV 480
Db 481 PGFHGTHCSSKVDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
QY 481 PGFHGTHCSSKVDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
Db 541 GGTCMNRVNSFCVCANGFRGKOCDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
QY 541 GGTCMNRVNSFCVCANGFRGKOCDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
Db 601 FSVAMPLVAVIAACVVFCKRKRKRAQEKDDAEARKONEQNAVATMHHNGSGVGVVALASA 660
QY 601 FSVAMPLVAVIAACVVFCKRKRKRAQEKDDAEARKONEQNAVATMHHNGSGVGVVALASA 660
Db 661 SLGGKTSNSGLTFDGGNPNIINKTWKSVNNICASAAAAAADADECLMTGGYVASV 720
QY 661 SLGGKTSNSGLTFDGGNPNIINKTWKSVNNICASAAAAAADADECLMTGGYVASV 720
Db 721 ADNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGSSAKGASGGGPGAEGKRISVL 780
QY 721 ADNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGSSAKGASGGGPGAEGKRISVL 780
Db 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGSAGTAQOORSVVCCTPHM 833
QY 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGSAGTAQOORSVVCCTPHM 833

RESULT 2 A31246 #type complete
ENTRY neurogenic protein Delta precursor - fruit fly (Drosophila)
TITLE melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
16-Feb-1997
ACCESSIONS A31246
REFERENCE Kopczynski, C.C.; Alton, A.K.; Fechtel, K.; Kooh, P.J.;
Muskhavitch, M.A.T.
#journal Genes Dev. (1988) 2:1723-1735
#title Delta, a Drosophila neurogenic gene, is transcriptionally
complex and encodes a protein related to blood coagulation
factors and epidermal growth factor of vertebrates.
#cross-references MUID:89196890
#accession A31246
#molecule_type mRNA
#residues 1-832 #label KOP
#cross-references GB:Y00222
GENETICS
#gene FlyBase:Dl
#cross-references FlyBase:FBgn0000463
SUMMARY #length 832 #molecular-weight 88943 #checksum 636
Query Match 99.1%; Score 6107; DB 2; Length 832;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 822; Conservative 8; Mismatches 2; Indels 1; Gaps 1;
Db 1 MWIKCLLTAFTCTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
QY 1 MWIKCLLTAFTCTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
Db 61 KXTRFRLCLKHQATIDTTSQCTYGVITPILGENSVNLTDAQRFQNKGFNPQFPF 120
QY 61 KXTRFRLCLKHQATIDTTSQCTYGVITPILGENSVNLTDAQRFQNKGFNPQFPF 120
Db 121 SWPFTGSLVIAEAWHDNNSGNARTNKLQLRLLVQVLEVSSEWKNKSESQTSLEYDF 180
```

```
QY 121 SWPFTGSLVIAEAWHDNNSGNARTNKLQLRLLVQVLEVSSEWKNKSESQTSLEYDF 180
Db 181 RVTCDLNYTSGCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPKCAKCEHGH 240
QY 181 RVTCDLNYTSGCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPKCAKCEHGH 240
Db 241 DRPNQCVQLGHWKALGALNCEVLEPNCIHGTCNKPWTICNEHGGGLYCNDLNYCTNHRP 300
QY 241 DRPNQCVQLGHWKALGALNCEVLEPNCIHGTCNKPWTICNEHGGGLYCNDLNYCTNHRP 300
Db 301 CKNGGTCFNTGEGLYCKCAPGYSGDDCENEIYSCDADVNPONGGTCIDEPHTKTYKC 360
QY 301 CKNGGTCFNTGEGLYCKCAPGYSGDDCENEIYSCDADVNPONGGTCIDEPHTKTYKC 360
Db 361 HCRNWSGKMKCEKVLTCSDKPCCHQICRNVRPGLSGKGGYQCECPICYSGPNCDLQLD 420
QY 361 HCRNWSGKMKCEKVLTCSDKPCCHQICRNVRPGLSGKGGYQCECPICYSGPNCDLQLD 420
Db 421 NCSPNPCIINGGSCQPSGKICIPAGSGTRCETNIDCLGHQCENGGTCIDMNVQYRCQCV 480
QY 421 NCSPNPCIINGGSCQPSGKICIPAGSGTRCETNIDCLGHQCENGGTCIDMNVQYRCQCV 480
Db 481 PGFHGTHCSSKVDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
QY 481 PGFHGTHCSSKVDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
Db 541 GGTCMNRVNSFCVCANGFRGKOCDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
QY 541 GGTCMNRVNSFCVCANGFRGKOCDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
Db 601 FSVAMPLVAVIAACVVFCKRKRKRAQEKDDAEARKONEQNAVATMHHNGSGVGVVALASA 660
QY 601 FSVAMPLVAVIAACVVFCKRKRKRAQEKDDAEARKONEQNAVATMHHNGSGVGVVALASA 660
Db 661 SMGGKTSNSGLTFDGGNPNIINKTWKSVNNICASAAAAAADADECLMTGGYVASV 720
QY 661 SLGGKTSNSGLTFDGGNPNIINKTWKSVNNICASAAAAAADADECLMTGGYVASV 720
Db 721 ADNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGTSAGKASGGGPGAEGKRISVL 780
QY 721 ADNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGTSAGKASGGGPGAEGKRISVL 780
Db 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGTD-GTAQOORSVVCCTPHM 832
QY 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGTAGTAQOORSVVCCTPHM 833

RESULT 3 S00670 #type complete
ENTRY gene Delta protein precursor - fruit fly (Drosophila)
TITLE melanogaster)
ALTERNATE_NAMES gene Dl protein
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
16-Feb-1997
ACCESSIONS S00670
REFERENCE S00670
#authors Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
#journal EMBO J. (1987) 6:3431-3440
#title The neurogenic gene Delta of Drosophila melanogaster is
expressed in neurogenic territories and encodes a putative
transmembrane protein with EGF-like repeats.
#accession S00670
#molecule_type mRNA
#residues 1-880 #label VAE
#cross-references EMBL:X06289
GENETICS
#gene Delta
#cross-references FlyBase:FBgn0000463
KEYWORDS transmembrane protein
FEATURE
1-18. #domain signal sequence #status predicted #label sig\
```

19-88 #product gene Delta protein #status predicted #label MAT
SUMMARY #length 880 #molecular-weight 94643 #checksum 7597

Query Match 98.1%; Score 6045; DB 2; Length 880;

Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 815; Conservative 9; Mismatches 7; Indels 1; Gaps 1;

```

Db 1 MHMNCILTAICFTVYIVQVSSGSFELRLKTFESNDHGDNGRCSSGSDATGKCLGS 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MHMNCILTAICFTVYIVQVSSGSFELRLKTFESNDHGDNGRCSSGSDATGKCLGS 60
Db 61 CKTFEPLCKHYQATIDTTSCTGYDVITPILGNSVNLDAQRONKGFPIQPFESF 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CKTFEPLCKHYQATIDTTSCTGYDVITPILGNSVNLDAQRONKGFPIQPFESF 120
Db 121 SWPQFSLIVEAMHDNTNNGARTKLLIQLLVQVLEVSSEMTKNSQSQYTSLEYDF 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 SWPQFSLIVEAMHDNTNNGARTKLLIQLLVQVLEVSSEMTKNSQSQYTSLEYDF 180
Db 181 RVTCDLNTYSGGCAKFCRPRDSDSGHSTCSETEGELICLTGMOGDYCHIRPKAKGCEHGC 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 RVTCDLNTYSGGCAKFCRPRDSDSGHSTCSETEGELICLTGMOGDYCHIRPKAKGCEHGC 240
Db 241 DKPACQVQLGKALGALNECVLEPCIHGTCKPMTICINEGWSGLYCNOIDLNTCTNHRP 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 DKPACQVQLGKALGALNECVLEPCIHGTCKPMTICINEGWSGLYCNOIDLNTCTNHRP 300
Db 301 CKNGTGFNTGEGLYTCKAPGSGDDCENEIYSCDADVPQNGSTCIDEPHTGTGKC 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 CKNGTGFNTGEGLYTCKAPGSGDDCENEIYSCDADVPQNGSTCIDEPHTGTGKC 360
Db 361 HCRNGWSKMEKEVLYTSDKPRCHOGICRNVRPGLSGGQGYQCCPPIGYSPPNDLQD 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 361 HCRNGWSKMEKEVLYTSDKPRCHOGICRNVRPGLSGGQGYQCCPPIGYSPPNDLQD 420
Db 421 NCSNPNCINGSCOPSGKICIPSGFRCETNIDCIGHOCENGSTCIDMAYNORCCV 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 421 NCSNPNCINGSCOPSGKICIPSGFRCETNIDCIGHOCENGSTCIDMAYNORCCV 480
Db 481 PGFHGTHSSRYDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSYDIDECSSGPGCHN 540
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 481 PGFHGTHSSRYDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSYDIDECSSGPGCHN 540
Db 541 GGTGMNRVNSPEVCANCFRQKQDEESYDTPAHQYGAITQARADGLANAQVYLAV 600
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 541 GGTGMNRVNSPEVCANCFRQKQDEESYDTPAHQYGAITQARADGLANAQVYLAV 600
Db 601 FSVAMPLVAVIAACVPCMKRRKRAOEKDNAEARKQDNQNAVATMHHNGSAVVALASA 660
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 601 FSVAMPLVAVIAACVPCMKRRKRAOEKDNAEARKQDNQNAVATMHHNGSAVVALASA 660
Db 661 SMGKTSNSGLTFDGNPNIIKNTWDRSVNNICASAAAAAADADECILMGGYVASV 720
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 661 SMGKTSNSGLTFDGNPNIIKNTWDRSVNNICASAAAAAADADECILMGGYVASV 720
Db 721 ADNNNANSDFCVAPLQRAKSKQKLTNPTLMDHRSFAGTSAKAGSGGPGAAEGKRISVL 780
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 721 ADNNNANSDFCVAPLQRAKSKQKLTNPTLMDHRSFAGTSAKAGSGGPGAAEGKRISVL 780
Db 781 GEGSYCSORWPSLAAGVAGDLFTOLMAAASVAGTD-CTAQQQRVYVCGTPH 831
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 781 GEGSYCSORWPSLAAGVAGDLFTOLMAAASVAGTD-CTAQQQRVYVCGTPH 831
```

RESULT 4
ENTRY 150719 #type complete
TITLE C-Delta-1 - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1996
ACCESSIONS 150719
REFERENCE 150719
#authors Henrique, D.; Adam, J.; Myat, A.; Chitnis, A.; Lewis, J.;

Ish-Horowitz, D.
Nature (1995) 375:787-790
Expression of a Delta homologue in prospective neurons in the chick.
#cross-references MUID:95319507
#accession 150719
#status preliminary; translated from GB/EMBL/DBJ
#molecule-type mRNA
#residues 1-728 #label HEN
#cross-references EMBL:U26590; NID:9882411; PID:9882412
SUMMARY #length 728 #molecular-weight 79861 #checksum 1765

Query Match 31.0%; Score 1910; DB 2; Length 728;
Best Local Similarity 50.1%; Pred. No. 0.00e+00;
Matches 263; Conservative 78; Mismatches 168; Indels 16; Gaps 15;

```

Db 20 QVDSGVFELKIQEYFNKGLLSNRCCGSGPGAGQOCCCKTFEPLCKHYQASVP 79
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 19 QVHSSGSFELRLKTFESNDHGRDNGRCSSGSDATGKCLGSKTRFRVCLKHQYATIDT 78
Db 80 EPPCTYGAITPVLGANSFVDPDAGADPAFNPRIPEPFTWPGTSLIETALHTDSP 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 79 TSQCTYGDVITPILGNSVNLDAQRONKGFPIQPFESWPQFSLIVEAMH-DTN 137
Db 140 DLTETENPERLSRLATQRIHLAVGEWSODLSSGTDLKYRFRVCDHHYGECSVRC 199
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 138 NSGNARTNKLIIQLLVQVLEVSSEMTKNSQSQYTSLEYDFRYVCDLNTYSGGCAKFC 197
Db 200 RPRDRFHFHTGGERKEKCNPMGKQGYCTEPICLPGCDEHGFCDKPECKCRVMOGR 259
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 198 RPRDSDSGHSTCSETEGELICLTGMOGDYCHIRPKAKGC-E-HGDKRPMQCVQLGKGA 255
Db 260 YDCECIRYPGLHGTQOQPMQCNCOEGWGLFCNODLNTCTHHKPKCKNGATCTNTGGSY 319
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 256 ICNECVLEPCIHGTCKPMTICINEGWSGLYCNOIDLNTCTNHRPCKNGTCFNTGEGLY 315
Db 320 TCSCRPGYTGSSCEIENNCDA--NPCKNGSGCTD-LE-NS-YSCTCPFGFKGKCELSA 374
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 316 TCKCAPGYSDDCENEIYSCDADVPQNGSTCIDEPHTGTGKCHCAGWSGKKEEYV 375
Db 375 MTCADGCPENGG-R-C-TD-NPDG-GYSCRCPLGYSGFKEKIDYCSSSPANGAQCVD 429
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 376 LTCSDPKPHOGICRNVRPGLSGGQGYQCCPPIGYSPPNDLQDNCSPNPIINGSC-Q 434
Db 430 LGNSYICQOQAGFTGRHCDNVDCAFPVNGTCTQDGVNDYSCTCPPEYGNKNCSTPV 489
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 435 -PSG-KCICPAGFSGTRCTETNIDCIGHOCENGSTCIDMAYNORCCVYGFHGTCSKV 492
Db 490 SRCEHNPCHNGATCHEHSNRYCECARGGLGNCQELPEPPOGR 534
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 493 DCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSYDIDECSSGP 537
```

RESULT 5
ENTRY 148324 #type complete
TITLE DELTA-like 1 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997
ACCESSIONS 148324
REFERENCE 148324
#authors Bettenhausen, B.; de Angelis, M.H.; Simon, D.; Guenet, J.L.; Gossler, A.
#journal Development (1995) 121:2407-2418
#title Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta.
#cross-references MUID:95401858
#accession 148324
#status preliminary; translated from GB/EMBL/DBJ
#molecule-type mRNA
#residues 1-722 #label RBS
#cross-references EMBL:X80903; NID:g806569; PID:g806570


```
|||||
QY 542 GTCMNRVNSFECVANGFRKQKDEESYDSTPDHAQYGAFTQARADGLTAQVLLAVE 601
Db 181 SVAMPLVAVTAACVFCMKR 200
QY 602 SVAMPLVAVTAACVFCMKR 621

RESULT 8
ENTRY A36666 #type complete
TITLE serrate protein precursor - fruit fly (Drosophila melanogaster)
ORGANISM #formal name Drosophila melanogaster
DATE 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 16-Feb-1997
ACCESSIONS A36666
REFERENCE A36666
#authors Fleming, R.J.; Scottgale, T.N.; Diederich, R.J.; Artavanis-Tsakonas, S.
#journal Genes Dev. (1990) 4:2188-2201
#title The gene Serrate encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in Drosophila melanogaster.
#cross-references M01D:91099666
#accession A36666
#status preliminary
#molecule_type mRNA
#residues 1-1404 #label FLE
GENETICS #cross-references GB:M35759
#gene FlyBase:Ser
#keywords #cross-references FlyBase:FBgn0004197
#transmembrane protein
#summary #length 1404 #molecular-weight 150245 #checksum 434

Query Match 20.5%; Score 1266; DB 2; Length 1404;
Best Local Similarity 34.8%; Pred. No. 9,53e-231;
Matches 207; Conservative 143; Mismatches 199; Indels 46; Gaps 35;

Db 63 CNLIALI-LILLVHKISAGNFELLEISNTNSHLNGYCCGPAELRATKTTGCSPT 121
QY 6 CULTAFCIFVIVQ-VHSSGSFELRLKYSNDHGRDNEGRCCSGSDGANGKCLG-S-CK 62
Db 122 TAFRLDLKEVQTEGASISGCSFGNATKTIAGSSFVSDP-----G-VGAIIVLPT 174
QY 63 TRRVVCLKHQAQID-TT-SQ-CYIGDVITPILIGENSVNLTDARFONKQFTNIDPFS 119
Db 175 FRTKSFETLILQ-LDMYNTSYDAERLIETSYGVILP-SPWKTLIDHIGRNARITR 232
QY 120 FSPHGFSLIVEAMHDNTNSGNARTKLLIQRLLVQVLEYSSWKTKNSQYTSLEYD 179
Db 233 VRVQCAVYVNTTCTTCRPRDQFGHYAGCSBQKLCNLNGMVGNCDEAICKAGCDPVH 292
QY 180 FRTCOLNLYTSGSCAFRCRDRDSFGHSTCSFGEIICLGMQGDYHIFKCAKGC--H 237
Db 293 GKCDRPECECRPMGRPLCNECMVTPGCKHSGCSAMKVCVCDTNMGILCDODLFCG 352
QY 238 GHCDKPNQCCQLGKMGKALCNECVLEPNCHIGTCN-KPWTICICNEGSGHLYCNODLNYCT 296
Db 353 THEPCRHGCGCENTAPDKYCTCAEGLISGQCEIVEHPC-A-TRPCANGGCTCTKTSNRT 410
QY 297 NHRPCRNNGGCFNTGELTYCKCAPGYSGDCENEIYSCDADVPNCPNGGTC-IDEPH-T 354
Db 411 QAQVYTSRHSRWMGRPVRRSSSMRSLDHLRPPGQALNGSSSGSLVLSGLQLOQOLAPD 470
QY 355 KTG-YK-CH-CAN-GWSGKCEE-KVLTG--SDRPHQIGICRNVRLGGS---KGQ-G-- 401
Db 471 FTCDCAAGWTGPTCEINIDECAGPCPEHGCTIDLIGFRCPECPPEWHGDCQVAVNECE 530
QY 402 YGCECIGTSGRNCIDLQDNCSNPNCLNGSSC-QP-SG-KCICPAGSGRCENIDDCI 458
Db 531 APHSAGIAANALLTTATATIGSNLSSTALLAALTSAVASTSLAIGCINAKECRNPGS 590
```

```
QY 459 GHQCCN-GGTCT--DMVQYR-CQ-CVPGFHTGHCSSKVDLCL-IRPCANGGTCILNLND 512
Db 591 FACICKEGMSGVTAENDDCV-GQCRNGATCIDLYNDRACASGFTGRCEFD 644
QY 513 YQCTCRAGFTGKDCSDVIDECCSGCHNGTGMNRVNSFECVANGFRKQKDEE 567

RESULT 9
ENTRY S16148 #type complete
TITLE gene serrate protein precursor - fruit fly (Drosophila melanogaster)
ORGANISM #formal name Drosophila melanogaster
DATE 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change 16-Feb-1997
ACCESSIONS S16148; S16878
REFERENCE S16148
#authors Thomas, U.; Speicher, S.A.; Knust, E.
#journal Development (1991) 111:749-761
#title The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs.
#cross-references M01D:91347903
#accession S16148
#molecule_type mRNA
#residues 1-1408 #label TH01
REFERENCE S16878
#authors Thomas, U.
#submission submitted to the EMBL Data Library, November 1990
#accession S16878
#molecule_type mRNA
#residues 1-1351,'T',1353-1408 #label TH02
GENETICS #cross-references EMBL:X56811
#gene Ser
#keywords #cross-references FlyBase:FBgn0004197
#classification #superfamily EGF homology
#keywords glycoprotein; transmembrane protein
FEATURE
1-84
85-1408
85-1221 #domain extracellular #status predicted #label EXT\
288-316 #domain EGF homology #label EG01\
319-348 #domain EGF homology #label EG02\
355-388 #domain EGF homology #label EG03\
395-488 #domain EGF homology #status atypical #label EG04\
495-526 #domain EGF homology #label EG05\
533-608 #domain EGF homology #status atypical #label EG06\
615-645 #domain EGF homology #label EG07\
652-683 #domain EGF homology #label EG08\
727-796 #domain EGF homology #label EG09\
803-834 #domain EGF homology #status atypical #label EG10\
841-876 #domain EGF homology #label EG11\
883-914 #domain EGF homology #label EG12\
921-952 #domain EGF homology #label EG13\
997-1060 #domain EGF homology #label EG14\
1222-1246 #region cysteine-rich\
1247-1408 #domain transmembrane #status predicted #label TM1\
132,196,247,331, #domain intracellular #status predicted #label INT\
412,452,558,739,
965,977,1004,1030,
1150
#binding site carbohydrate (Asn) (covalent) #status predicted
SUMMARY #length 1408 #molecular-weight 150645 #checksum 5466

Query Match 20.5%; Score 1266; DB 2; Length 1408;
Best Local Similarity 34.8%; Pred. No. 9,53e-231;
Matches 207; Conservative 143; Mismatches 199; Indels 46; Gaps 35;

Db 67 CNLIALI-LILLVHKISAGNFELLEISNTNSHLNGYCCGPAELRATKTTGCSPT 125
```

```
QY 6 CLLTAFICFTVIVQ-VHSSGSFELRLKYPFNDHGRDNEGRCCSGSDGATGKCLG-S-CK 62
Db 126 TAFRLCLKEYQTEOGASISTCSCFNATFKILGSSSFVLSDP-----G-VGAIVLPT 178
QY 63 TRFRVCLKHQYQID-TT-SQ-CTYGDVITPILGNSVNLTAQRFQNGFTNPQFPFS 119
Db 179 FRWTKSFLLILOA-LDMYNTSPDAERLLEETSYSGVILP-SPWKTLDHIGNARITYR 236
QY 120 FSWPQTFSLIVEAWHTNNSGNARTNKKLIQRLLOVQVLEVSSEWTKNSQYTSLEYD 179
Db 237 VRVOCAYTYNTCTTFCPRDDQFQHYACGSEGOIKLCLNGWQVNCBAICAKAGCDPVH 296
QY 180 FRVTCDLNLYGSCAKFCRPRDDSGHSTCSETGEIICLTGWQDYCHIPKAKGCE--H 237
Db 297 GKCDRPGCECRWGRPLNCEMYVPGCKHGSCGSAWKVCVDTNWGGILCDDQDLNFCG 356
QY 238 GHCDRPNQCVCQKLGKALGALNCEVLEPNCIHGTCN-KPWTICINEGWGLYCNQDLNYCT 296
Db 357 THEPCKHGGTCENTAPDKYKRCACBGLSGEOCEIYEHPC-A-TRPCRNGGTCTLTKSNT 414
QY 297 NHRPCKNGGTCTENTGEGLYTCKAPYSGDDCENEIYSCDADVNPQNGGTC-IDEPH-T 354
Db 415 OAQVYRTSHGRSNMGRPVRRSSMSRLDHLRPEGOALNGSSSLVSLGSLQLOQLAPD 474
QY 355 KTG-YK-CH-CAN-GWSGKMCCE-KVLTG--SDKPCHQICRNVRPLGS---KGQ-G-- 401
Db 475 FTCDCAAGWTGPTCEINIDECAGGCEHGGTCIDLIGFRCPCPPWHGVDQVQVNECE 534
QY 402 YOCEPIGYSGPNCIDLQDNCSPNCGGSC-QP-SG-KCICPAGFSGTRCETNIDDC 458
Db 535 APHSAGIAANALLTTATAIIGSNLSSTALLAALTSAAVASTSLAIGPCINAKERNQPGS 594
QY 459 GHOCEN-GGTCTI--DMVNYQR-CQ-CVPFGHGTCHSSKVDLCL-IRPCANGGTCLNND 512
Db 595 FACICKEGSGVTCANLDDCV-GQCRNGATCIDLVNDYRCACASGTRGDCETD 648
QY 513 YOCTCRAGFTGKDCSDVIDECSSGPHNGGTGTCMNRVNSFECVANGFRKQCDDEE 567

RESULT 10
ENTRY #type complete
TITLE Transmembrane protein precursor - zebra fish
ORGANISM #formal_name Brachydanio rerio #common_name zebra fish
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 12-Sep-1997
ACCESSIONS S42612
REFERENCE S42612
#authors Bierkamp, C.; Campos-Ortega, J.A.
#journal Mech. Dev. (1993) 43:87-100
#title A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern of transcription during early embryogenesis.
#accession S42612
#status preliminary
#molecule_type mRNA
#residues 1-2437 #label BIE
#cross-references EMBL:X69088; NID:9433866; PID:9433867
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin repeat homology
FEATURE
1915-1947 #domain ankyrin repeat homology #label AN1\
1948-1980 #domain ankyrin repeat homology #label AN2\
1982-2014 #domain ankyrin repeat homology #label AN3\
2048-2080 #domain ankyrin repeat homology #label AN5
SUMMARY #length 2437 #molecular-weight 262306 #checksum 4021
Query Match 16.0%; Score 987; DB 2; Length 2437;
Best Local Similarity 39.1%; Pred. No. 4.40e-172;
Matches 154; Conservative 82; Mismatches 124; Indels 34; Gaps 23;
Db 643 CRRKPCDYKCIDKINGECYCEPGYSGSMCNINIDDCALNPNCHNGGTCLDGVNSTCLC 702
QY 197 CPRDDSFHGSCTSEFG-EIICLTGWQGDYCHIP--KCA-KGCEHG-HC-DKPNQ--CVC 248
```

```
Db 703 PDGFRDATCLSOHNESSNP-CIHGSCLDQINSYRCVCEAGWGRNCDININECLSN-PC 760
QY 249 QLGKWKALC-----NECVLEPNCIHGTC-N--KPWTICINEGWGLYCNQDLNCTHRRPC 301
Db 761 VLGGTCKDWTSG-YLCTCRAGFSGPNCOMNINEC-AS-NPCLNOGSCIDD--V-AGFKCN 814
QY 302 KNGGTCTFNTGEGLYTCKAPYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTGKCH 361
Db 815 CMLPYTGEVCEN-VLAPCSPPRCKNGVCRE-SEDF-----QSFSCNCPAGWQGTCEVDI 868
QY 362 CANGWSGKMCERKVLTCSDKPCHQG-ICRNVRPLGSKGQGYQCPCPIGYSNCDLQL 419
Db 869 NECVRNPNCNGVCENLRGFGQRCNPGTGCALCENIDDDCEPNPCSNGGVQODRVNGFV 928
QY 420 DNGSPNPGGSCQ--PSG-KCICPAGFSGTRCETNIDDCDHLGHQCEGTCIDMYNOYR 476
Db 929 CVCLAGFRERCAEDIDECVAPCRNGGNTDCVNSYTCSPAGFGGICNEINTPCTES 988
QY 477 CQCVPGFHGTHCSKVDLCLIRPCANGGTCLNNDYQCTCRAGFTGKDCSDVIDECSSG 536
Db 989 SCFNGGTCVDIGTSSFCVCLPGLFTGNYCQHDVNE 1022
QY 537 PCHNGGTGTCMNRVNSFECVANGFRGKQCDDESYD 570

RESULT 11
ENTRY #type fragment
TITLE Motch B protein - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 31-Oct-1997
ACCESSIONS A49175; PH1570; S32113
REFERENCE A49175
#authors Lardelli, M.; Lendahl, U.
#journal Exp. Cell Res. (1993) 204:364-372
#title Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety of tissues.
#cross-references MUID:93178563
#accession A49175
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-1203 #label LAR
#cross-references EMBL:X68279; NID:927989; PID:9287990
#experimental_source embryo
#note sequence extracted from NCBI backbone (NCBIP:126158)
COMMENT This protein has many EGF repeats and lin-12/Notch repeats.
COMMENT This protein is one of the neurogenic proteins controlling the decision between ectodermal and neural fate for cells in the early embryo.
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin repeat homology
SUMMARY #length 1203 #checksum 910
Query Match 15.7%; Score 970; DB 2; Length 1203;
Best Local Similarity 40.9%; Pred. No. 1.57e-168;
Matches 146; Conservative 74; Mismatches 108; Indels 29; Gaps 17;
Db 401 YSQVNECLSNPCITHGNTGCLSGYKCLDAGWGVNCEYDKNECLSNP-CQNGGTCLN 459
QY 225 YCHIPKC-ARGCBEHGHG-DK-PN-QVCQLGWKALC-----NECVLEPNCIHG-TCNK-- 273
Db 460 NGYRCCKKGFYGNQVNDICASN-PCLNQGTCTFDDVSG-YTCHCMLPYTKGNQTVL 517
QY 274 -PWTICINEGWGLYCNQDLNCTNHRPCANGGTCTFEGGLYTCRCAPYSGDDCENEI 332
Db 518 APCSP--NPPCENAAVCKEAPNFES-FSCLCAPGWQGRKCTVDYDECIKPCMNNGVCHNT 574
QY 333 YSCDADVNPQNGGTCTIDEPTTKTGKCHCANGWSGKMCERKVLTCSDKPC-HQICRN 391
Db 575 Q--GS-----YVCECPGFSGMDCEEDINDCLANPCONGSCVDHVNFTFSCQHPGFIGD 627
QY 197 CPRDDSFHGSCTSEFG-EIICLTGWQGDYCHIP--KCA-KGCEHG-HC-DKPNQ--CVC 248
```

```

QY 392 RPLGSGKGGYOCCEPCIGYSGPNCLQJLNCSPNFCINGGSC--QPSG-KCICPAGFSGT 448
Db 628 KCOTDMNECISEPCKNGTSDYVNSTCTCPAGHGVHCENIDECESSCFNGTCVD 687
QY 449 RCEFNIDDCJGHOCENGTCDYVNYRCQCPGFGHGHSSKGLCLIRPCANSGTCLN 508
Db 688 GINSFCLCEVGFPGFPCILHDINECSSNPCLNAGTCVGLGTYRCICLGYTGKNCQ 744
QY 509 LNDYOCCTCAAGFTGKCKSCSYDIDECSSGPCNGGICANRVNSFEVCANGFRGKCD 565

RESULT 12
ENTRY A49128 #type complete
TITLE cell-fate determining gene Notch2 protein - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 12-Sep-1997

ACCESSIONS A49128
REFERENCE A49128
#authors Weinmaster, G.; Roberts, V.J.; Lemke, G.
#journal Development (1992) 116:931-941
#title Notch2: a second mammalian Notch gene.
#cross-references MUID:93202015
#accession A49128
#status preliminary: not compared with conceptual translation
#molecule_type mRNA
#residues 1-2471 #label WEI
#experimental_source Schwann cell
#note sequence extracted from NCBI backbone (NCBIP:127811)
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin
repeat homology
FEATURE
1876-1908 #domain ankyrin repeat homology #label AN1\
1909-1941 #domain ankyrin repeat homology #label AN2\
1943-1975 #domain ankyrin repeat homology #label AN3\
2009-2041 #domain ankyrin repeat homology #label AN5
SUMMARY #length 2471 #molecular_weight 265367 #checksum 5929

Query Match 15.7%; Score 967; DB 2; Length 2471;
Best Local Similarity 44.1%; Pred. No. 6,656-168;
Matches 139; Conservative 64; Mismatches 90; Indels 22; Gaps 16;

```

```

DATE purple urchin
13-May-1992 #sequence_revision 17-Sep-1997 #text_change 20-Mar-1998
ACCESSIONS A40136; B40136; C40136; A29316; A43131
REFERENCE A40136
#authors Delgado-Illio-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff, R.A.
#journal J. Mol. Evol. (1989) 29:314-327
#title Structural analysis of the uEGF gene in the sea urchin Strongylocentrotus purpuratus reveals more similarity to vertebrate than to invertebrate genes with EGF-like repeats.
#cross-references MUID:90112459
#accession A40136
#status preliminary
#molecule_type mRNA
#residues 1-114 #label DEL
#cross-references GB:X17530; NID:g10225; PID:g667061
#accession B40136
#status preliminary: not compared with conceptual translation
#molecule_type DNA
#residues 181-251,329-370,'R',372-408,'RA',411-441 #label DE2
#accession C40136
#status preliminary: not compared with conceptual translation
#molecule_type DNA
#residues A29316
#authors Hursh, D.A.; Andrews, M.E.; Raff, R.A.
#journal Science (1987) 237:1487-1490
#title A sea urchin gene encodes a polypeptide homologous to epidermal growth factor.
#cross-references MUID:87319677
#accession A29316
#status preliminary
#molecule_type mRNA
#residues 'S',280-481,786-1064 #label HUR
#cross-references GB:M17421; NID:g161474; PID:g552260
REFERENCE A43131
#authors Hunt, L.T.; Barker, W.C.
#journal FASEB J. (1989) 3:1760-1764
#title Avidin-like domain in an epidermal growth factor homolog from a sea urchin.
#cross-references MUID:89196806
#contents annotation
#note EGF homology repeats 10-17 are spliced out in the short form (fibropellin ID).
CLASSIFICATION #superfamily C1r/C1s repeat homology; EGF homology
FEATURE
1-19 #domain signal sequence #status predicted #label SIG\
20-1064 #product fibropellin I #status predicted #label FIB\
20-56 #domain EGF homology #label EGF1\
57-175 #domain C1r/C1s repeat homology #label CSR\
176-213 #domain EGF homology #label EGF2\
214-251 #domain EGF homology #label EGF3\
252-289 #domain EGF homology #label EGF4\
290-327 #domain EGF homology #label EGF5\
328-365 #domain EGF homology #label EGF6\
366-403 #domain EGF homology #label EGF7\
404-441 #domain EGF homology #label EGF8\
442-479 #domain EGF homology #label EGF9\
480-517 #domain EGF homology #label EGF10\
518-555 #domain EGF homology #label EGF11\
556-593 #domain EGF homology #label EGF12\
594-631 #domain EGF homology #label EGF13\
632-669 #domain EGF homology #label EGF14\
670-707 #domain EGF homology #label EGF15\
708-745 #domain EGF homology #label EGF16\
746-783 #domain EGF homology #label EGF17\
784-821 #domain EGF homology #label EGF18\
822-859 #domain EGF homology #label EGF19\
860-897 #domain EGF homology #label EGF20\
898-935 #domain EGF homology #label EGF21\
936-1064 #region avidin-like\

```


2258-2264, 'V', 2266-2406, 'R', 2408-2444, 'L', 2446-2703
##label WHA
the authors translated the codon ATC for residue 49 as
Thr, ATT for residue 2044 as Arg, GTA for residue 2265
as Ala, CGC for residue 2407 as His, and CTT for
residue 2445 as Arg

GENETICS

notch

```
##Cross-references FlyBase:FBgn0004647
53/3 84/3 171/3 240/3 3
```

| INTRONS | EXTRACATON |
|---------------------------------------------------------|---------------------------------------------------------|
| 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3 | 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3 |

CLASSIFICATION #superfamily notch protein; ankyrin repeat homology; EGF homology

homology

KEYWORDS differentiation; membrane protein; tandem repeat

FEATURE-
1950-

| | #domain | ankyrin repeat | homology | #label | AN1\ |
|-----------|---------|----------------|----------|--------|------|
| 1950-1982 | | | | | |
| 1983-2015 | | | | | |

| 1983-2013 | #domain | ankyrin repeat | homology | #label | AN2\ |
|-----------|---------|----------------|----------|--------|------|
| 2017-2049 | #domain | ankyrin repeat | homology | #label | AN3\ |

| Accession | Protein | Label |
|-----------|--------------------------------|-------|
| 2017-2045 | ankyrin repeat homology | ANK5 |
| 2050-2082 | domain ankyrin repeat homology | ANK4 |

| Accession | Protein | Label | AN5 |
|-----------|----------------|-------|-----|
| 2050-2062 | ankyrin repeat | #1 | AN4 |
| 2083-2115 | ankyrin repeat | #1 | AN5 |

| NAME | FUNCTION | DESCRIPTION | DATE |
|-------|----------|-----------------------------------------------|------|
| 2000 | 2110 | #length 2703 #molecular-weight 288876 #checks | |
| MMARY | | | |

[illegible]

Best Local

Matches 148; Conservative 83; Mismatches 116; Indels 3

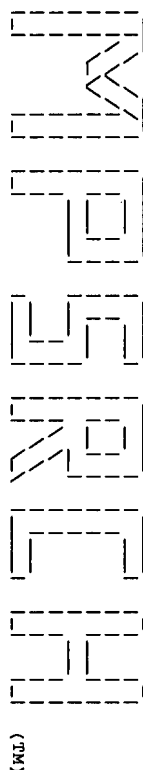
| DATE | DESCRIPTION | AMOUNT | BALANCE |
|------------|-----------------|--------|---------|
| 1950-01-01 | OPENING BALANCE | 100.00 | 100.00 |
| 1950-01-15 | PAYROLL | 50.00 | 50.00 |
| 1950-02-01 | RENT | 25.00 | 25.00 |
| 1950-02-15 | UTILITIES | 10.00 | 15.00 |
| 1950-03-01 | SALES | 150.00 | 165.00 |
| 1950-03-15 | PAYROLL | 50.00 | 115.00 |
| 1950-04-01 | RENT | 25.00 | 90.00 |
| 1950-04-15 | UTILITIES | 10.00 | 80.00 |
| 1950-05-01 | SALES | 120.00 | 200.00 |
| 1950-05-15 | PAYROLL | 50.00 | 150.00 |
| 1950-06-01 | RENT | 25.00 | 125.00 |
| 1950-06-15 | UTILITIES | 10.00 | 115.00 |
| 1950-07-01 | SALES | 180.00 | 295.00 |
| 1950-07-15 | PAYROLL | 50.00 | 245.00 |
| 1950-08-01 | RENT | 25.00 | 220.00 |
| 1950-08-15 | UTILITIES | 10.00 | 210.00 |
| 1950-09-01 | SALES | 140.00 | 350.00 |
| 1950-09-15 | PAYROLL | 50.00 | 300.00 |
| 1950-10-01 | RENT | 25.00 | 275.00 |
| 1950-10-15 | UTILITIES | 10.00 | 265.00 |
| 1950-11-01 | SALES | 160.00 | 425.00 |
| 1950-11-15 | PAYROLL | 50.00 | 375.00 |
| 1950-12-01 | RENT | 25.00 | 350.00 |
| 1950-12-15 | UTILITIES | 10.00 | 340.00 |
| 1951-01-01 | CLOSING BALANCE | | 340.00 |

```

Dh 488 NINCEISPCNNESSCLDDPCTFCVCMFGTGTQCEIDIDECOSNPLCNDGCHING 547
Qy 227 HIPC-KAGCEH-GHC-DKPN-QCVQLGKALCN-ECLV-EPN-CIH-GIC-NK--P 274
Dh 548 FKCSALGFTGARCQIINDDQSO-PCRNIGICHSIAG-YSGCEPBGYTGTSCIEINID 605
Qy 275 WTCICNEBGMGLYCNOPLNCTNHRPCKNGGTCEPNTGEGLYTCKCAGAYSDDCEMEIYS 334
Dh 606 CDS--NQCNRG-KRIDD--VNS-FKCLCDBGYTYITQKQINCESNPQCFDGHCD-R 657
Qy 335 CDAVNPQONQGTIDBPHRTGTGKCHCANGWSKMCDEEVLTCSDPCH-QGICRWVR 393
Dh 658 --VGS----YYCQCAAGTSGKNCENVECHSNPCNNATCITDGINSYKCGVGFQHC 712
Qy 394 GLSGKGOGTQCECPIYISGPRCDLQJLNCSPNFCINNGSC-QP-SG-KCICPAGFSGTRC 450
Dh 713 EKNVDECISSPCANNVCIDQVNGYKCECPRGFYDAHCLSDVDECASNPCVNEGRCEGI 772
Qy 451 ETNIDDCIGHQCEHNGGVCIDMVNGYRQCVPGFHGTHCSSKVDCLIRPCANGSTCLNUN 510
Dh 773 NEFICHCPRGYTGKRCCLDIDECSSNQCQHGTYIDKLANFSCQCMNGYTGQKCEFNID 832
Qy 511 NDYOCYCRAGTGTGDCSDVIDDECSSGCHNGGTCIMNRYVNSFEVCYANGFGRKQCDSEYD 570
Dh 833 CVT--NPGNGGTCLIDKYN 850
Qy 571 SVYEDAHQYCATTOARRDG 589

```

Search completed: Fri Jun 12 11:53:23 1998
Job time : 156 secs.


 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Jun 12 11:53:42 1998; MasPar time 21.43 Seconds
 974.872 Million cell updates/sec
 Tabular output not generated.

Title: >US-08-083-590A-2
 Description: (1-833) from US08083590A.pep (1 of 2)
 Perfect Score: 6164
 Sequence: 1 MHWIKCLNFAFCFTYIVQV.....GSGAGTAQQRSVYCGTPHM 833

Scoring table:
 PAM 150
 Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot35
 1:swiss1

Statistics: Mean 50.001; Variance 79.902; scale 0.626

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------------|-------------------------|-----------|
| 1 | 6045 | 98.1 | 880 | DL_DROME | NEUROGENIC LOCUS DELTA | 0.00e+00 |
| 2 | 1826 | 29.6 | 722 | DL1_MOUSE | DELTA-LIKE PROTEIN 1 P | 0.00e+00 |
| 3 | 1819 | 29.5 | 714 | DL1_RAT | DELTA-LIKE PROTEIN 1 P | 0.00e+00 |
| 4 | 1266 | 20.5 | 1408 | SERR_DROME | SERRATE PROTEIN PRECUR | 5.51e-285 |
| 5 | 987 | 16.0 | 2437 | FBP1_MOUSE | NEUROGENIC LOCUS NOTCH | 1.98e-212 |
| 6 | 959 | 15.6 | 1064 | FBP1_STRPU | FIBROBLASTIN 1 PRECURSO | 3.37e-205 |
| 7 | 956 | 15.5 | 2524 | NOTC_XENLA | NEUROGENIC LOCUS NOTCH | 2.01e-204 |
| 8 | 951 | 15.4 | 2703 | NOTC_DROME | NEUROGENIC LOCUS NOTCH | 3.91e-203 |
| 9 | 928 | 15.1 | 2531 | NTC1_MOUSE | NEUROGENIC LOCUS NOTCH | 3.32e-197 |
| 10 | 920 | 14.9 | 2531 | NTC1_RAT | NEUROGENIC LOCUS NOTCH | 1.43e-192 |
| 11 | 910 | 14.8 | 2444 | NTC1_HUMAN | NEUROGENIC LOCUS NOTCH | 1.75e-189 |
| 12 | 898 | 14.6 | 2318 | NTC3_MOUSE | FIBROBLASTIN C PRECURSO | 7.43e-185 |
| 13 | 881 | 14.3 | 570 | FBP3_STRPU | NEUROGENIC LOCUS NOTCH | 4.10e-185 |
| 14 | 798 | 12.9 | 1964 | NTC4_MOUSE | NEUROGENIC LOCUS NOTCH | 1.93e-155 |
| 15 | 765 | 12.4 | 2133 | CRB_DROME | CRUMBS PROTEIN PRECURS | 2.48e-127 |
| 16 | 654 | 10.6 | 383 | DLK_HUMAN | DELTA-LIKE PROTEIN PRE | 2.48e-127 |
| 17 | 654 | 10.6 | 385 | DLK_MOUSE | DELTA-LIKE PROTEIN PRE | 2.48e-127 |
| 18 | 647 | 10.5 | 1480 | SLIT_DROME | SLIT PROTEIN PRECURSOR | 4.62e-124 |
| 19 | 641 | 10.4 | 1429 | SLIT2_MOUSE | SLIT-2 PROTEIN PRECURS | 4.62e-124 |
| 20 | 593 | 8.8 | 1295 | GLP1_CAEL | GLP-1 PROTEIN PRECURSO | 1.33e-99 |
| 21 | 493 | 8.0 | 2199 | TEN1_HUMAN | TENASCIN PRECURSOR (TN | 2.78e-87 |
| 22 | 489 | 7.9 | 515 | APX1_CAEL | APX-1 PROTEIN PRECURSO | 2.66e-86 |
| 23 | 464 | 7.5 | 1808 | TEN1_CHICK | TENASCIN PRECURSOR (TN | 3.40e-80 |

| RESULT | 1 | STANDARD | PRT | 880 AA. |
|--------|-----------------------------------------------------------------------|----------|-----|---------|
| ID | DL_DROME | | | |
| AC | P10041 | | | |
| DT | 01-MAR-1989 (REL. 10, CREATED) | | | |
| DT | 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE) | | | |
| DT | 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE) | | | |
| DE | NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR. | | | |
| GN | DL. | | | |
| OS | DROSOPHILA MELANOGASTER (FRUIT FLY). | | | |
| OC | EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | VAESSIN H., BREMER K.A., KNUST E., CAMPOS-ORTEGA J.A.; | | | |
| RL | EMBO J. 6:3431-3440(1987). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 422-621 FROM N.A. | | | |
| RX | MEDLINE: 87218537 | | | |
| RA | KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D., VAESSIN H., | | | |
| RA | CAMPOS-ORTEGA J.A.; | | | |
| RL | EMBO J. 6:761-766(1987). | | | |
| RN | [3] | | | |
| RP | PATTERN OF TRANSCRIPTION. | | | |
| RX | MEDLINE: 91209246 | | | |
| RA | HENDLIN M., KRAMATSCHEK B., CAMPOS-ORTEGA J.A.; | | | |
| RL | DEVELOPMENT 110:905-914(1990). | | | |
| CC | - FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL | | | |
| CC | IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL | | | |
| CC | CELL LINEAGES. | | | |
| CC | - SUPRACELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | |
| CC | - SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART | | | |
| CC | OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, | | | |
| CC | THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES. | | | |
| CC | - NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL, IT IS | | | |
| CC | CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPLETE | | | |
| CC | FOR BINDING WITH THE NOTCH PROTEIN. | | | |
| CC | - SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS. | | | |
| CC | - SIMILARITY: TO DROSOPHILA SERRATE PROTEIN. | | | |
| DR | EMBL: X06289; G7853; - | | | |
| DR | EMBL: X05140; G929563; - | | | |
| DR | PIR: S00670; S00670. | | | |
| DR | PIR: A26637; A26637. | | | |
| DR | HSSP: P00725; ICBB. | | | |
| DR | FLYBASE: FBgn0000463; DL. | | | |
| DR | PROSITE: PS00010; ASX_HYDROXYL; 3. | | | |
| DR | PROSITE: PS00022; EGF_1; 9. | | | |
| DR | PROSITE: PS01186; EGF_2; 9. | | | |
| DR | PROSITE: PS01187; EGF_CA; 2. | | | |

ALIGNMENTS

KW DIFFERENTIATION; NEUROGENESIS; REPEAT; TRANSMEMBRANE;
KW EGF-LIKE DOMAIN; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 18
FT CHAIN 19 880
FT DOMAIN 19 653
FT TRANSMEM 654 677
FT DOMAIN 678 880
FT DOMAIN 227 258
FT DOMAIN 256 289
FT DOMAIN 291 329
FT DOMAIN 331 372
FT DOMAIN 374 416
FT DOMAIN 418 451
FT DOMAIN 453 489
FT DOMAIN 491 527
FT DOMAIN 529 565
FT DISULFID 231 240
FT DISULFID 235 246
FT DISULFID 248 257
FT DISULFID 260 271
FT DISULFID 266 277
FT DISULFID 279 288
FT DISULFID 295 307
FT DISULFID 301 317
FT DISULFID 319 328
FT DISULFID 335 348
FT DISULFID 342 360
FT DISULFID 362 371
FT DISULFID 378 388
FT DISULFID 383 404
FT DISULFID 406 415
FT DISULFID 422 433
FT DISULFID 427 439
FT DISULFID 441 450
FT DISULFID 457 468
FT DISULFID 462 477
FT DISULFID 479 488
FT DISULFID 495 506
FT DISULFID 500 515
FT DISULFID 517 526
FT DISULFID 533 544
FT DISULFID 538 553
FT DISULFID 555 564
FT CARBOHYD 98 98
FT CARBOHYD 137 137
FT CARBOHYD 167 167
FT CARBOHYD 649 649
FT CONFLICT 437 438
FT CONFLICT 459 459
FT CONFLICT 490 490
SQ SEQUENCE 880 AA; 94643 MW; E967E562 CRC32;

Query Match 98.1%; Score 6045; DB 1; Length 880;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 815; Conservative 9; Mismatches 7; Indels 1; Gaps 1;

Db 1 MHWINCLLTAFCFTVIQVHSSGSEFELRLKYSNDHGRNDRGRCSDGATGKCLGS 60
QY 1 MHWINCLLTAFCFTVIQVHSSGSEFELRLKYSNDHGRNDRGRCSDGATGKCLGS 60
Db 61 KTRFRCLKHYQATIDTTSQCTYGDVITPILGENSVNLTDQRFQNKFTNPIQPPFSF 120
QY 61 KTRFRCLKHYQATIDTTSQCTYGDVITPILGENSVNLTDQRFQNKFTNPIQPPFSF 120
Db 121 SWPFTSLIWEAHDNTNNGARTNKLTLQRLVQVLEVSSEWTKNSQYTSLEYDF 180
QY 121 SWPFTSLIWEAHDNTNNGARTNKLTLQRLVQVLEVSSEWTKNSQYTSLEYDF 180
Db 181 RVTCDLNYGSGCAKFCPRDDSFHSTCSETGEIICLTGWQGDYCHIPKCAKGEHHC 240
QY 181 RVTCDLNYGSGCAKFCPRDDSFHSTCSETGEIICLTGWQGDYCHIPKCAKGEHHC 240
Db 241 DKPNQVCQLGKNGALCNCEVLEPNCIHGTCNKPTWCICNEGWGGLYCNDLNYCTNHRP 300

QY 241 DKPNQVCQLGKNGALCNCEVLEPNCIHGTCNKPTWCICNEGWGGLYCNDLNYCTNHRP 300
Db 301 CKNGTCTENTGEGLYTCKAPGYSDDCENEIYSCADVNPQNGGTCTIDEPHTTGYKC 360
QY 301 CKNGTCTENTGEGLYTCKAPGYSDDCENEIYSCADVNPQNGGTCTIDEPHTTGYKC 360
Db 361 HCRNGWSGKMECEKVLCTCDKPCDHOGICRNVRPGLSGGQGYQCECPITGYSGPNCDLQLD 420
QY 361 HCRNGWSGKMECEKVLCTCDKPCDHOGICRNVRPGLSGGQGYQCECPITGYSGPNCDLQLD 420
Db 421 NCSNPNCINGGSCQPSGKICPISGFSGRCTENIIDDCLGHQCENGGTCTIDMWNQYRCQCV 480
QY 421 NCSNPNCINGGSCQPSGKICPISGFSGRCTENIIDDCLGHQCENGGTCTIDMWNQYRCQCV 480
Db 481 PGFHGTHGSSKVDLCLIRPCANGGTCLMINDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
QY 481 PGFHGTHGSSKVDLCLIRPCANGGTCLMINDYQCTCRAGFTGKDCSVDIDECSSGPCHN 540
Db 541 GGTGMNRVNSPECVANGFRGKQCEDESVDSTFDAHGYGATTQARADGLANAQVVLIAV 600
QY 541 GGTGMNRVNSPECVANGFRGKQCEDESVDSTFDAHGYGATTQARADGLANAQVVLIAV 600
Db 601 FSVAMPLVAVIAACVVFVFCMKRRKRAQEKDNEARKQEQNAVATMHNGSAGVALASA 660
QY 601 FSVAMPLVAVIAACVVFVFCMKRRKRAQEKDNEARKQEQNAVATMHNGSAGVALASA 660
Db 661 SMGGTGSNSGLTFDGGNPNIIKNTWDSVNNICASAAAAAABDECLMYGYVASV 720
QY 661 SMGGTGSNSGLTFDGGNPNIIKNTWDSVNNICASAAAAAABDECLMYGYVASV 720
Db 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGTSKAGSGGGGPAEGKRISVL 780
QY 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGTSKAGSGGGGPAEGKRISVL 780
Db 781 GEGSYCSQRPWSLAAGVAGDLFIOLMAAASVAGTD-GTAQOORSVVCCTPH 831
QY 781 GEGSYCSQRPWSLAAGVAGDLFIOLMAAASVAGTD-GTAQOORSVVCCTPH 831
RESULT 2
ID DLL1_MOUSE STANDARD; PRT: 722 AA.
AC O61483;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA1).
GN DLL1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C X C57BL/6; TISSUE=EMBRYO;
RX MEDLINE; 95401858.
RA BETTENHAUSEN B., DE ANGELIS M.H., SIMON D., GUENET J.-L., GOSSLER A.;
RL DEVELOPMENT 121:2407-2418(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED IN THE PARAXIAL
MESODERM AND NERVOUS SYSTEM. EXPRESSED AT HIGH LEVELS IN ADULT
CC HEART AND AT LOWER LEVELS, IN ADULT LUNG.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED UNTIL DAY 15 IN THE EMBRYO.
CC EXPRESSION THEN DECREASES AND INCREASES AGAIN IN THE ADULT.
CC -!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: TO DROSOPHILA DELTA PROTEIN.
DR EMBL; X80903; G806570; -.
DR MGD; MGI:104659; DLL1.
DR PROSITE; PS00010; ASX HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 8.

FT DISULFID 883 894 BY SIMILARITY.
 FT DISULFID 888 903 BY SIMILARITY.
 FT DISULFID 905 914 BY SIMILARITY.
 FT DISULFID 921 932 BY SIMILARITY.
 FT DISULFID 926 941 BY SIMILARITY.
 FT DISULFID 943 952 BY SIMILARITY.
 FT CARBOHYD 152 152 POTENTIAL.
 FT CARBOHYD 156 156 POTENTIAL.
 FT CARBOHYD 196 196 POTENTIAL.
 FT CARBOHYD 247 247 POTENTIAL.
 FT CARBOHYD 331 331 POTENTIAL.
 FT CARBOHYD 412 412 POTENTIAL.
 FT CARBOHYD 452 452 POTENTIAL.
 FT CARBOHYD 558 558 POTENTIAL.
 FT CARBOHYD 739 739 POTENTIAL.
 FT CARBOHYD 965 965 POTENTIAL.
 FT CARBOHYD 977 977 POTENTIAL.
 FT CARBOHYD 1004 1004 POTENTIAL.
 FT CARBOHYD 1030 1030 POTENTIAL.
 FT CARBOHYD 1150 1150 POTENTIAL.
 FT CONFLICT 14 17 MISSING (IN REF. 2).
 FT CONFLICT 27 27 P -> A (IN REF. 2).
 FT CONFLICT 1352 1352 T -> S (IN REF. 2).
 SQ SEQUENCE 1408 AA; 150660 MW; A494A358 CRC32;

Query Match 20.5%; Score 1266; DB 1; Length 1408;
 Best Local Similarity 34.8%; Pred. No. 5,51e-285;
 Matches 207; Conservative 143; Mismatches 199; Indels 46; Gaps 35;

DB 67 CMLTALI-LILLVHKSIAAGNELEIEISNTSHLNGCGMPAELRAKTIQCSPT 125
 QY 6 CLTIAFCIFYVQ-VHSSGFELRLKRFSDHGDHNGRCCSGSDATGKCLG-S-CK 62
 DB 126 TAPRLCKEYOTTEOGASISTGCSFGNATYKILGSSSVLSDP-----G-VGALVLPPT 178
 QY 63 TRFVCKKHQOATID-TT-SQ-CITYGVITPILGSESVNLTDAPQFQNKGFNPQFPFS 119
 DB 179 FRMTKSTLLIOA-LDMYNTSYPAERLIETSYSGVLP-SPEKTLDHIGNARIYR 236
 QY 120 FSWGPTSLIVEAHWDHNNNSNARTNKLRLVQVLEVSSEKTKKSSQYTSLEYD 179
 DB 237 VAVOCATYNTCTTCFPRPDQFGHYACSEGOIKLCLNMGVNCBEALCKACADPVH 296
 QY 180 FRVYCDLNTYGGSCAKCRPRDSDFGHSTGSETEIILTMQDGYCHIRPCACGE--H 237
 DB 297 GKCDRECECECRPGWRPLCNECMVYPCQKHGSCNGSAMKVCYDITNMGILCDODLFCG 356
 QY 238 GHCDKPMQCVQQLGMKALCNEVLEPNCIHGTEN-KPMTICINMGWGLCNODLANCT 296
 DB 357 THEPCKRGICENTAPDKYRGTCAEGLSGECCIEVHRC-A-TRPCRNNGTCTLKTSMRT 414
 QY 297 NHRPCKNGGTCFNTGEGLYTCKCAPGYSGDCENEIYSCDADVNPCCONGTIC-IDEPT-T 354
 DB 415 QAOVYRSHGRNMRPRVRRSSMRSLDHLRPEGOALNGSSSGVLSGLQLOQOLAPD 474
 QY 355 KTG-YK-CH-CAN-GWSGKCEE-KVLTC--SDKPCRHGDIRNVRPGLS---KGQ-G-- 401
 DB 475 FTDCDCAAGWTPTCEINIDECAGPCEHGTICIDLGFRCECPPEMHGVCQVDVNECE 534
 QY 402 YQCECPIGYSGPNCDDQLDNCSPNCPINGGSC-QP-SG-KICPAGFGTCEITNIDCL 458
 DB 535 APSAGIAANALTTTATIGLSMLSTALALTAASVASTLSALIGPINAKECNOGGS 594
 QY 459 GHOCEN-GGTICI--DMVNOYR-CQ-CVPGFHGTICSSVVDLCT-IRPCANGTICINLND 512
 DB 595 FACICKRGWGTCAENLDCV-GGCRNGATCIDLVNDYRCACASGFTGRDCEIT 648
 QY 513 YQTCRAGFTGKDCVDIDBCSSPCHNGTICMKNVNSFECVCAENGFKQCDDEE 567

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
 GN NOTCH.
 OS BRACHYDANTIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
 OC EURARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
 OC OSTEIFRTHES; ACTINOPTERYGII; CYPRINIFORMES.
 NC [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO.
 RX MEDLINE; 94128602.
 RA BIERKAMP C., CAMPOS-ORTEGA J.A.;
 RL MECH. DEV. 43:87-100(1993).
 CC -1- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
 CC EMBRYO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
 CC NEURAL PLATE, NOTCHORD AND BRAIN VESICLES.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
 CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED,
 CC ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND
 CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
 CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
 CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
 CC BRAIN AND HEAD REGIONS.
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 DR EMBL; X69088; G433867; -.
 DR PROSITE; PS0010; ASX_HYDROXYL; 23.
 DR PROSITE; PS00022; EGF_1; 34.
 DR PROSITE; PS01186; EGF_2; 28.
 DR PROSITE; PS01187; EGF_CA; 22.
 KW DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
 KW TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
 FT SIGNAL 1 20
 FT CHAIN 21 2437
 FT DOMAIN 21 1724
 FT TRANSMEM 1725 1747
 FT DOMAIN 1748 2437
 FT DOMAIN 21 57
 FT DOMAIN 58 98
 FT DOMAIN 101 138
 FT DOMAIN 139 175
 FT DOMAIN 177 215
 FT DOMAIN 217 254
 FT DOMAIN 256 292
 FT DOMAIN 294 332
 FT DOMAIN 334 370
 FT DOMAIN 371 409
 FT DOMAIN 411 449
 FT DOMAIN 451 487
 FT DOMAIN 489 524
 FT DOMAIN 526 562
 FT DOMAIN 564 599
 FT DOMAIN 601 639
 FT DOMAIN 639 674
 FT DOMAIN 676 712
 FT DOMAIN 714 749
 FT DOMAIN 751 787
 FT DOMAIN 789 825
 FT DOMAIN 827 865
 FT DOMAIN 867 903
 FT DOMAIN 905 941
 FT DOMAIN 943 979
 FT DOMAIN 981 1017
 FT DOMAIN 1019 1055
 FT DOMAIN 1057 1093
 FT DOMAIN 1095 1143
 FT DOMAIN 1143 1179
 FT DOMAIN 1181 1217
 FT DOMAIN 1219 1263
 FT DOMAIN 1265 1303
 POTENTIAL.
 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 EGF-LIKE 1.
 EGF-LIKE 2.
 EGF-LIKE 3.
 EGF-LIKE 4.
 EGF-LIKE 5.
 EGF-LIKE 6.
 EGF-LIKE 7.
 EGF-LIKE 8.
 EGF-LIKE 9.
 EGF-LIKE 10.
 EGF-LIKE 11.
 EGF-LIKE 12.
 EGF-LIKE 13.
 EGF-LIKE 14.
 EGF-LIKE 15.
 EGF-LIKE 16.
 EGF-LIKE 17.
 EGF-LIKE 18.
 EGF-LIKE 19.
 EGF-LIKE 20.
 EGF-LIKE 21.
 EGF-LIKE 22.
 EGF-LIKE 23.
 EGF-LIKE 24.
 EGF-LIKE 25.
 EGF-LIKE 26.
 EGF-LIKE 27.
 EGF-LIKE 28.
 EGF-LIKE 29.
 EGF-LIKE 30.
 EGF-LIKE 31.
 EGF-LIKE 32.
 EGF-LIKE 33.

```
FT DOMAIN 1305 1344 EGF-LIKE 34. 775 760 FT DISULFID 775 BY SIMILARITY.
FT DOMAIN 1346 1382 EGF-LIKE 35. 786 777 FT DISULFID 786 BY SIMILARITY.
FT DOMAIN 1382 1423 EGF-LIKE 36. 804 793 FT DISULFID 804 BY SIMILARITY.
FT DOMAIN 1423 1461 3 X LIN/NOTCH REPEATS. 813 798 FT DISULFID 813 BY SIMILARITY.
FT REPEAT 1461 1486 LIN/NOTCH 1. 824 824 FT DISULFID 824 BY SIMILARITY.
FT REPEAT 1486 1520 LIN/NOTCH 2. 831 842 FT DISULFID 831 BY SIMILARITY.
FT REPEAT 1520 1561 LIN/NOTCH 3. 853 853 FT DISULFID 853 BY SIMILARITY.
FT REPEAT 1561 1621 6 X ANK MOTIF REPEATS. 864 855 FT DISULFID 864 BY SIMILARITY.
FT REPEAT 1621 1691 ANK MOTIF 1. 882 871 FT DISULFID 882 BY SIMILARITY.
FT REPEAT 1691 1740 ANK MOTIF 1. 891 876 FT DISULFID 891 BY SIMILARITY.
FT REPEAT 1740 1974 ANK MOTIF 1. 902 893 FT DISULFID 902 BY SIMILARITY.
FT REPEAT 1974 2007 ANK MOTIF 1. 920 909 FT DISULFID 920 BY SIMILARITY.
FT REPEAT 2007 2040 ANK MOTIF 1. 929 914 FT DISULFID 929 BY SIMILARITY.
FT REPEAT 2040 2074 ANK MOTIF 1. 940 931 FT DISULFID 940 BY SIMILARITY.
FT DOMAIN 2074 2276 POLY-GLN (OPA-REPEAT). 958 947 FT DISULFID 958 BY SIMILARITY.
FT DISULFID 2276 25 35 BY SIMILARITY. 967 952 FT DISULFID 967 BY SIMILARITY.
FT DISULFID 25 35 BY SIMILARITY. 978 969 FT DISULFID 978 BY SIMILARITY.
FT DISULFID 35 45 BY SIMILARITY. 1034 1023 FT DISULFID 1034 BY SIMILARITY.
FT DISULFID 45 56 BY SIMILARITY. 1043 1028 FT DISULFID 1043 BY SIMILARITY.
FT DISULFID 56 73 BY SIMILARITY. 1054 1045 FT DISULFID 1054 BY SIMILARITY.
FT DISULFID 73 86 BY SIMILARITY. 1072 1061 FT DISULFID 1072 BY SIMILARITY.
FT DISULFID 86 97 BY SIMILARITY. 1081 1066 FT DISULFID 1081 BY SIMILARITY.
FT DISULFID 97 116 BY SIMILARITY. 1092 1083 FT DISULFID 1092 BY SIMILARITY.
FT DISULFID 116 126 BY SIMILARITY. 1120 1099 FT DISULFID 1120 BY SIMILARITY.
FT DISULFID 126 137 BY SIMILARITY. 1129 1114 FT DISULFID 1129 BY SIMILARITY.
FT DISULFID 137 154 BY SIMILARITY. 1140 1131 FT DISULFID 1140 BY SIMILARITY.
FT DISULFID 154 163 BY SIMILARITY. 1158 1147 FT DISULFID 1158 BY SIMILARITY.
FT DISULFID 163 174 BY SIMILARITY. 1167 1152 FT DISULFID 1167 BY SIMILARITY.
FT DISULFID 174 194 BY SIMILARITY. 1178 1169 FT DISULFID 1178 BY SIMILARITY.
FT DISULFID 194 203 BY SIMILARITY. 1196 1185 FT DISULFID 1196 BY SIMILARITY.
FT DISULFID 203 214 BY SIMILARITY. 1205 1190 FT DISULFID 1205 BY SIMILARITY.
FT DISULFID 214 232 BY SIMILARITY. 1216 1207 FT DISULFID 1216 BY SIMILARITY.
FT DISULFID 232 242 BY SIMILARITY. 1242 1223 FT DISULFID 1242 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY. 1251 1236 FT DISULFID 1251 BY SIMILARITY.
FT DISULFID 253 271 BY SIMILARITY. 1262 1253 FT DISULFID 1262 BY SIMILARITY.
FT DISULFID 271 280 BY SIMILARITY. 1282 1269 FT DISULFID 1282 BY SIMILARITY.
FT DISULFID 280 291 BY SIMILARITY. 1291 1274 FT DISULFID 1291 BY SIMILARITY.
FT DISULFID 291 311 BY SIMILARITY. 1302 1302 FT DISULFID 1302 BY SIMILARITY.
FT DISULFID 311 320 BY SIMILARITY. 1320 1320 FT DISULFID 1320 BY SIMILARITY.
FT DISULFID 320 331 BY SIMILARITY. 1332 1332 FT DISULFID 1332 BY SIMILARITY.
FT DISULFID 331 349 BY SIMILARITY. 1361 1350 FT DISULFID 1361 BY SIMILARITY.
FT DISULFID 349 358 BY SIMILARITY. 1370 1370 FT DISULFID 1370 BY SIMILARITY.
FT DISULFID 358 369 BY SIMILARITY. 1381 1381 FT DISULFID 1381 BY SIMILARITY.
FT DISULFID 369 375 BY SIMILARITY. 1400 1400 FT DISULFID 1400 BY SIMILARITY.
FT DISULFID 375 380 BY SIMILARITY. 1411 1411 FT DISULFID 1411 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY. 1422 1422 FT DISULFID 1422 BY SIMILARITY.
FT DISULFID 397 408 BY SIMILARITY. 957 957 FT CARBOHYD 957 POTENTIAL.
FT DISULFID 408 415 BY SIMILARITY. 1177 1177 FT CARBOHYD 1177 POTENTIAL.
FT DISULFID 415 428 BY SIMILARITY. 1487 1487 FT CARBOHYD 1487 POTENTIAL.
FT DISULFID 428 437 BY SIMILARITY. 1585 1585 FT CARBOHYD 1585 POTENTIAL.
FT DISULFID 437 448 BY SIMILARITY.
FT DISULFID 448 455 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 466 475 BY SIMILARITY.
FT DISULFID 475 486 BY SIMILARITY.
FT DISULFID 486 493 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 503 512 BY SIMILARITY.
FT DISULFID 512 523 BY SIMILARITY.
FT DISULFID 523 541 BY SIMILARITY.
FT DISULFID 541 550 BY SIMILARITY.
FT DISULFID 550 561 BY SIMILARITY.
FT DISULFID 561 578 BY SIMILARITY.
FT DISULFID 578 587 BY SIMILARITY.
FT DISULFID 587 598 BY SIMILARITY.
FT DISULFID 598 605 BY SIMILARITY.
FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 616 625 BY SIMILARITY.
FT DISULFID 625 636 BY SIMILARITY.
FT DISULFID 636 643 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 653 662 BY SIMILARITY.
FT DISULFID 662 673 BY SIMILARITY.
FT DISULFID 673 691 BY SIMILARITY.
FT DISULFID 691 700 BY SIMILARITY.
FT DISULFID 700 711 BY SIMILARITY.
FT DISULFID 711 728 BY SIMILARITY.
FT DISULFID 728 737 BY SIMILARITY.
FT DISULFID 737 748 BY SIMILARITY.
FT DISULFID 748 755 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.

Query Match 16.08; Score 987; DB 1; Length 2437;
Best Local Similarity 39.1%; Pred. No. 1.98e-212;
Matches 154; Conservative 82; Mismatches 124; Indels 34; Gaps 23;

Db 643 CKRRPCDYGKIDKINGECVCEPCEPGSGSMCNINIDCALNPNCHNGTICDGVNSFTCLC 702
QY 197 CRPRDSDSGHSTCSEYG-EIICLTGWQGDYCHP--KCA-KGCEHG-HC-DKPNQ--CVC 248
Db 703 PDGFRDATCISQHNCESSNP-CHGSCLDQINSYRCVCEAGWGRNCDININECLSN-PC 760
QY 249 QLGWKALC-----NECVLEPNCIHGTC-N--KPWTCICNEGWGLYCNQDLNYCTNHRPC 301
Db 761 VNGGTCADMFTSG-YLCTCRAGFSGPCNQMINEC-AS-NECLNQGSCIDD--V-AGEKCN 814
QY 302 KNGGTCFTGEGLYTCKCAPYSGDCENETISCDADNPQNGGTCIDSPHTKTYKCH 361
Db 815 CMLPYTGEVCEN-VLAPCSPRCKNKGVCRE-SEDF----QSFSCNCPAGWQGTCEVDI 868
QY 362 CANGWSKMCERKVLTCSDKPCHQG-ICRNVRLGSKGGYQCECPIGYSGPNCIDLQ 419

Note: remainder of annotations omitted.
```


| | | | | |
|----|----------|------|------|----------------|
| FT | DISULFID | 150 | 165 | BY SIMILARITY. |
| FT | DISULFID | 167 | 176 | BY SIMILARITY. |
| FT | DISULFID | 183 | 194 | BY SIMILARITY. |
| FT | DISULFID | 188 | 203 | BY SIMILARITY. |
| FT | DISULFID | 205 | 214 | BY SIMILARITY. |
| FT | DISULFID | 221 | 232 | BY SIMILARITY. |
| FT | DISULFID | 226 | 242 | BY SIMILARITY. |
| FT | DISULFID | 244 | 253 | BY SIMILARITY. |
| FT | DISULFID | 260 | 271 | BY SIMILARITY. |
| FT | DISULFID | 265 | 280 | BY SIMILARITY. |
| FT | DISULFID | 282 | 291 | BY SIMILARITY. |
| FT | DISULFID | 298 | 311 | BY SIMILARITY. |
| FT | DISULFID | 305 | 320 | BY SIMILARITY. |
| FT | DISULFID | 322 | 331 | BY SIMILARITY. |
| FT | DISULFID | 338 | 349 | BY SIMILARITY. |
| FT | DISULFID | 343 | 358 | BY SIMILARITY. |
| FT | DISULFID | 360 | 369 | BY SIMILARITY. |
| FT | DISULFID | 375 | 386 | BY SIMILARITY. |
| FT | DISULFID | 380 | 397 | BY SIMILARITY. |
| FT | DISULFID | 399 | 408 | BY SIMILARITY. |
| FT | DISULFID | 415 | 428 | BY SIMILARITY. |
| FT | DISULFID | 422 | 437 | BY SIMILARITY. |
| FT | DISULFID | 439 | 448 | BY SIMILARITY. |
| FT | DISULFID | 455 | 466 | BY SIMILARITY. |
| FT | DISULFID | 460 | 475 | BY SIMILARITY. |
| FT | DISULFID | 477 | 486 | BY SIMILARITY. |
| FT | DISULFID | 493 | 504 | BY SIMILARITY. |
| FT | DISULFID | 498 | 513 | BY SIMILARITY. |
| FT | DISULFID | 515 | 524 | BY SIMILARITY. |
| FT | DISULFID | 531 | 542 | BY SIMILARITY. |
| FT | DISULFID | 536 | 551 | BY SIMILARITY. |
| FT | DISULFID | 553 | 562 | BY SIMILARITY. |
| FT | DISULFID | 569 | 579 | BY SIMILARITY. |
| FT | DISULFID | 574 | 588 | BY SIMILARITY. |
| FT | DISULFID | 590 | 599 | BY SIMILARITY. |
| FT | DISULFID | 606 | 617 | BY SIMILARITY. |
| FT | DISULFID | 611 | 626 | BY SIMILARITY. |
| FT | DISULFID | 628 | 637 | BY SIMILARITY. |
| FT | DISULFID | 644 | 654 | BY SIMILARITY. |
| FT | DISULFID | 649 | 663 | BY SIMILARITY. |
| FT | DISULFID | 665 | 674 | BY SIMILARITY. |
| FT | DISULFID | 681 | 692 | BY SIMILARITY. |
| FT | DISULFID | 686 | 701 | BY SIMILARITY. |
| FT | DISULFID | 703 | 712 | BY SIMILARITY. |
| FT | DISULFID | 719 | 729 | BY SIMILARITY. |
| FT | DISULFID | 724 | 738 | BY SIMILARITY. |
| FT | DISULFID | 740 | 749 | BY SIMILARITY. |
| FT | DISULFID | 756 | 767 | BY SIMILARITY. |
| FT | DISULFID | 761 | 776 | BY SIMILARITY. |
| FT | DISULFID | 778 | 787 | BY SIMILARITY. |
| FT | DISULFID | 794 | 805 | BY SIMILARITY. |
| FT | DISULFID | 799 | 814 | BY SIMILARITY. |
| FT | DISULFID | 816 | 825 | BY SIMILARITY. |
| FT | DISULFID | 832 | 843 | BY SIMILARITY. |
| FT | DISULFID | 837 | 854 | BY SIMILARITY. |
| FT | DISULFID | 856 | 865 | BY SIMILARITY. |
| FT | DISULFID | 872 | 883 | BY SIMILARITY. |
| FT | DISULFID | 877 | 892 | BY SIMILARITY. |
| FT | DISULFID | 894 | 903 | BY SIMILARITY. |
| FT | DISULFID | 910 | 921 | BY SIMILARITY. |
| FT | DISULFID | 915 | 930 | BY SIMILARITY. |
| FT | DISULFID | 932 | 941 | BY SIMILARITY. |
| FT | DISULFID | 986 | 997 | BY SIMILARITY. |
| FT | DISULFID | 991 | 1006 | BY SIMILARITY. |
| FT | DISULFID | 1008 | 1017 | BY SIMILARITY. |
| FT | DISULFID | 1024 | 1035 | BY SIMILARITY. |
| FT | DISULFID | 1029 | 1044 | BY SIMILARITY. |
| FT | DISULFID | 1046 | 1055 | BY SIMILARITY. |
| FT | DISULFID | 1062 | 1073 | BY SIMILARITY. |
| FT | DISULFID | 1067 | 1082 | BY SIMILARITY. |
| FT | DISULFID | 1084 | 1093 | BY SIMILARITY. |
| FT | DISULFID | 1100 | 1121 | BY SIMILARITY. |
| FT | DISULFID | 1115 | 1130 | BY SIMILARITY. |

| | | | | |
|----|----------|------|--------|---------------------|
| FT | DISULFID | 1132 | 1141 | BY SIMILARITY. |
| FT | DISULFID | 1148 | 1159 | BY SIMILARITY. |
| FT | DISULFID | 1153 | 1168 | BY SIMILARITY. |
| FT | DISULFID | 1170 | 1179 | BY SIMILARITY. |
| FT | DISULFID | 1186 | 1197 | BY SIMILARITY. |
| FT | DISULFID | 1191 | 1206 | BY SIMILARITY. |
| FT | DISULFID | 1208 | 1217 | BY SIMILARITY. |
| FT | DISULFID | 1224 | 1243 | BY SIMILARITY. |
| FT | DISULFID | 1237 | 1252 | BY SIMILARITY. |
| FT | DISULFID | 1254 | 1263 | BY SIMILARITY. |
| FT | DISULFID | 1270 | 1283 | BY SIMILARITY. |
| FT | DISULFID | 1275 | 1292 | BY SIMILARITY. |
| FT | DISULFID | 1294 | 1303 | BY SIMILARITY. |
| FT | DISULFID | 1310 | 1321 | BY SIMILARITY. |
| FT | DISULFID | 1315 | 1333 | BY SIMILARITY. |
| FT | DISULFID | 1335 | 1344 | BY SIMILARITY. |
| FT | DISULFID | 1351 | 1362 | BY SIMILARITY. |
| FT | DISULFID | 1356 | 1371 | BY SIMILARITY. |
| FT | DISULFID | 1373 | 1382 | BY SIMILARITY. |
| FT | DISULFID | 1390 | 1401 | BY SIMILARITY. |
| FT | DISULFID | 1395 | 1412 | BY SIMILARITY. |
| FT | DISULFID | 1414 | 1423 | BY SIMILARITY. |
| FT | CARBOHYD | 462 | 462 | POTENTIAL. |
| FT | CARBOHYD | 887 | 887 | POTENTIAL. |
| FT | CARBOHYD | 958 | 958 | POTENTIAL. |
| FT | CARBOHYD | 1178 | 1178 | POTENTIAL. |
| FT | CARBOHYD | 1487 | 1487 | POTENTIAL. |
| FT | CARBOHYD | 1508 | 1508 | POTENTIAL. |
| FT | CARBOHYD | 1584 | 1584 | POTENTIAL. |
| SO | SEQUENCE | 2524 | 275123 | MM: E6358481 CRC32; |

Query Match 15.5%; Score 956; DB 1; Length 2524;
Best Local Similarity 36.0%; Pred. No. 2.01e-204;
Matches 134; Conservative 96; Mismatches 112; Indels 30; Gaps 24;

| | | | |
|----|-----|--------------------------------------------------------------|-----|
| Db | 221 | CNPSPLNGCTGRQFD-DTSYDCTCLPFGSQNCENIDCPSPNCRNGTCVGVNFTYN | 279 |
| Qy | 226 | CHTPKRAKS--CEHGCHDPRNOCVQGLKMGKALCNCEVLE-P--NCHG-TC-N--KPMT | 276 |
| Db | 280 | COCPPDWTQYCTEDVDECQLMPNACNGSGTCNTYGG--YNCVGVNGWTGEDCSENIDDC | 338 |
| Qy | 277 | CICNEGSGGLYCNQDNLNCT-NHRPCKNGGTGFNTGEGLYTCKCAGSGGDDCENETYS | 335 |
| Db | 339 | -ANA-ACHSGATC-HD-RASFT-CECPHGRGTGLCHLDN-ACISNPKCBGSGNC-DTNP- | 390 |
| Qy | 336 | DADVNCQNGGTCIDEPHRTGTGKCHCANGWSKMEERYVLCTCSDPRCHQGI-CRNV | 394 |
| Db | 391 | VNGKA--I-CTCPPGYTGPACNNDVDECSIGANPCHEGRCRTWTLSFGCNCPOGYAGPR | 447 |
| Qy | 395 | LGSKGQGYCECPDIGSGFNCIDLQDNCG--PWCINGSCOPS-G--KCICPAGFSGR | 449 |
| Db | 448 | CEIDVNECLSNPCONDSTGLDQIGEPQICMPEBGLYCEINIDECASNPCILHNGKCIDK | 507 |
| Qy | 450 | CEINIDDCGHQCENGCTCIDMVNVRCCVGFPHGTHCSKVDLLIRPCANGGTCNLW | 509 |
| Db | 508 | INEFRCPTFGSGNLQHDPECTSTPCKNGAKCLDGENSTYCOCTBEFTGRHCEQDIN | 567 |
| Qy | 510 | NNDYOCCTCRAGFTGKDCSVIDECSSGPHNGGTCMNRVNSECVANGFRRGKOCESBY | 569 |
| Db | 568 | ECIP-DPCHYGT | 578 |
| Qy | 570 | DVTEFDAHQYA | 581 |

RESULT 8
ID NOTC_DROME STANDARD; PRT: 2703 AA.
AC P07207; P04154;
DE 01-NOV-1986 (REL. 03, CREATED)
DE 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.
GN N.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).


```
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 616 BY SIMILARITY.
FT DISULFID 611 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 654 BY SIMILARITY.
FT DISULFID 648 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 686 701 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 730 BY SIMILARITY.
FT DISULFID 724 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 893 BY SIMILARITY.
FT DISULFID 895 904 BY SIMILARITY.
FT DISULFID 911 923 BY SIMILARITY.
FT DISULFID 917 932 BY SIMILARITY.
FT DISULFID 934 943 BY SIMILARITY.
FT DISULFID 950 961 BY SIMILARITY.
FT DISULFID 955 970 BY SIMILARITY.
FT DISULFID 972 981 BY SIMILARITY.
FT DISULFID 988 999 BY SIMILARITY.
FT DISULFID 993 1008 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1026 1037 BY SIMILARITY.
FT DISULFID 1031 1046 BY SIMILARITY.
FT DISULFID 1048 1057 BY SIMILARITY.
FT DISULFID 1064 1075 BY SIMILARITY.
FT DISULFID 1069 1084 BY SIMILARITY.
FT DISULFID 1086 1095 BY SIMILARITY.
FT DISULFID 1102 1113 BY SIMILARITY.
FT DISULFID 1107 1122 BY SIMILARITY.
FT DISULFID 1124 1133 BY SIMILARITY.
FT DISULFID 1140 1160 BY SIMILARITY.
FT DISULFID 1155 1169 BY SIMILARITY.
FT DISULFID 1171 1180 BY SIMILARITY.
Note: remainder of annotations omitted.
```

```
Query Match 15.4%; Score 951; DB 1; Length 2703;
Best Local Similarity 39.1%; Pred. No. 3,916-203;
Matches 148; Conservative 83; Mismatches 116; Indels 32; Gaps 24;

Db 488 NINCESHPQNESSCLDDPGTFRCVCMPTGTGCELDIDECOSNPLNDSTCHDKING 547
QY 227 HTPKC-ANGCH-GHC-DKPN--QCVCOLGWMGALCN-ECVL-EPN-CIH-GTC-NK--P 274
Db 548 FKSCALGFTGARCOINIDDCOSQ-PCRNRGICHDISNG-YSCBPPGYTGSCININD 605
QY 275 WTCICNEGWSGLXCNODLNTNHRPCNKGCTCFNTGSLVTKCKAPYSBGDCNEIYS 334
Db 606 CDS--NPHRG-KCID--VNS-EKCLDPEYTGICQKQINECSNFCQEDGHCOD-R- 657
QY 335 CDADVNPQONGGTCIDEPHTKTKGKCHCANGSMGKMEKVLTCSDKPKCH--QGICRNVBP 393
Db 658 -VGS-----YYQCOAGTSGKNCVAVNCHSNPCNNGATCIDGINSYCCQCVPGTGOHC 712
QY 394 GLGSKGOGYOCEPIGYSGPCDLDNCSNPNCLNGSC-QP-SG-KCICPAGSGTRC 450
Db 713 EKNVDECISSPCANNGVYIDOVNGKCECPGFYDAHCLTDVDECASNPCVNEGRCEDGI 772
QY 451 ETNIDDLGHCENGCTCIDMNVNORCQVPGFHTGCHSCSKVDLCLIRPCANGTCLNIN 510
```

```
Db 773 NEFICHPGYTKRCRCELIDECSSNPCOHGCTCYDKLNAFSCOCMPGYGQKCEINID 832
QY 511 NDYOCTCRAGFTGKNCOSYDIDECSSGPCNNGSTCMNRNVSFCVCANGRGNQCEBESD 570
Db 833 CVT-NPCNGGTCIDKING 850
QY 571 SVTPDAHNGATTQARADG 589

RESULT 9
ID NTCL_MOUSE STANDARD; PRT; 2531 AA.
AC 001705;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE: 93194170.
RA FRANCO DEL AMO F., GENDRON-MAGUIRE M., SWIATEK P.J., JENKINS N.A.,
RL COPELAND N.G., GRIDLEY T.;
[2]
RN GENOMICS 15:259-264(1993).
RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE: 93048835.
RA FRANCO DEL AMO F., SMITH D.E., SWIATEK P.J., GENDRON-MAGUIRE M.,
RL GREENSPAN R.J., MCMAHON A.P., GRIDLEY T.;
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
DR EMBL: 211886; G288503;
DR MGD: MGI:97363; NOTCH1.
DR PROSITE: PS00010; ASX_HYDROXYL. 22.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 27.
DR PROSITE: PS01187; EGF_CA; 21.
DR DIFFERENTIATION: NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
KW TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT DOMAIN 19 1725
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 24 1425
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT DOMAIN 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT DOMAIN 1865 2075
FT REPEAT 1865 1910
FT REPEAT 1862 1912
FT REPEAT 1912 1942
FT REPEAT 1944 1975
FT REPEAT 1978 2009
FT REPEAT 2011 2042
FT REPEAT 2044 2075
FT CARBOHYD 888 888
FT CARBOHYD 959 959
FT CARBOHYD 1179 1179
FT CARBOHYD 1241 1241
FT CARBOHYD 1489 1489
FT CARBOHYD 1587 1587
FT SEQUENCE 2531 AA; 271312 MW; AD71189B CRC32;
```

| | | | | |
|-----------------------|--------|---------------------------------------------------------------|-------|-------------------------------------|
| Query Match | 15.19; | Score 928; | DB 1; | Length 2531; |
| Best Local Similarity | 38.7%; | Pred. No. 3.32e-197; | | |
| Matches | 147; | Conservative | 82; | Mismatches 113; Indels 38; Gaps 26; |
| Db | 465 | ATCLDQIGFQCICMPGYEGVYCEINTDECASSPCLHNGCHMDKIHFFCQCPCPKGFNGHL | 524 | |
| QY | 207 | STC--SETGEI--ICLTGWGDYCHI-PK-CAKG-CHEH-GHC-DK-PN-QCVCOLGKWGAL | 256 | |
| Db | 525 | CQYDVDCASTP-CCKNAKCLDGPNVTVCYCTGGYTGTCHVEDIDSC-DPDPCHYG-SC- | 580 | |
| QY | 257 | C----NECVLEPNCIHGT-C-NRP--WTC1CNBGGWGLYCQNQLNTCTNRPCKNCGTGF | 308 | |
| Db | 581 | KDGVATFTCLCPQYTGHHCTINETECHSQ--PCRHGTC-QD-RDNS-YLCLCLKGTTG | 635 | |
| QY | 309 | NTGEGLYCTKCAPYSGDDCENEIYSADVNPCQNGGTCTIDEPTHKTGYKCHCANGWSG | 368 | |
| Db | 636 | PNCEINLDCCASNPDCSGTC-----LD-KIDGEACACEPGYTGSMCNVNDICAGSPCH | 688 | |
| QY | 369 | KMCEEKVLTSCKDPHQOGICRNVRRPLGSKGQGQYCEPIGYSGPNCDLQLDNCSPNCI | 428 | |
| Db | 689 | NGGTCEGDIAGTCRCPEGYHDPTCLSEVEKSNPCIHG-ACRDGLNGYKCDCAFQWSG | 747 | |
| QY | 429 | NGGSCQP--SG-KVICPAGSFGTRCTNIDDLGHGCENGGTCTDMVNQRVCVPFFHG | 485 | |
| Db | 748 | TNCDINNNECESNPCVNGVGTKDKMTSGYVCTCREGSGPNQCNTNINECASNPCLNGTCTI | 807 | |
| QY | 486 | THCSKVLDLCIRPCANGGTCLNLNDYQCTCRAGFTGKDCSVDEICSSGPCNHNGTGM | 545 | |
| Db | 808 | DDVAGYKCNCPLPYTGATCE | 827 | |
| QY | 546 | NRVNSFEVCVANGFRGRQCD | 565 | |

RESULT 10

| ID | NTCL_RAT | STANDARD; | PRT; | 2531 AA. |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------|-----------|------|-------------------------------------------|
| AC | Q07008; | | | |
| DT | 01-NOV-1995 (REL. 32, CREATED) | | | |
| DT | 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE) | | | |
| DT | 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE) | | | |
| DE | NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR. | | | |
| GN | NOTCH1. | | | |
| OS | RATTUS NORVEGICUS (RAT). | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; | | | |
| OC | EUTHERIA; RODENTIA. | | | |
| [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE-SCHWANN CELL. | | | |
| RX | MEDLINE; 92111383. | | | |
| RA | WEINMASTER G.; ROBERTS V.J.; LEMKE G.; | | | |
| RL | DEVELOPMENT 113:195-205(1991). | | | |
| CC | -1- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER OF TISSUES. | | | |
| CC | -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | |
| CC | -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE ADULT. | | | |
| CC | -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS. | | | |
| CC | -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS. | | | |
| CC | -1- SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT. | | | |
| CC | -1- SIMILARITY: CONTAINS 6 ANK REPEATS. | | | |
| DR | ENBL; X57405; G57635; .. | | | |
| DR | PROSITE; PS00010; ASX-HYDROXYL; 22. | | | |
| DR | PROSITE; PS00022; EGF_1; 35. | | | |
| DR | PROSITE; PS01186; EGF_2; 26. | | | |
| DR | PROSITE; PS01187; EGF_CA; 21. | | | |
| DR | PROSITE; PS50039; FORK HEAD_3; 1. | | | |
| KW | DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN; TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN. | | | |
| KW | SIGNAL | | | |
| FT | CHAIN | 1 | 18 | POTENTIAL. |
| FT | DOMAIN | 19 | 2531 | NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1. |
| FT | DOMAIN | 19 | 1723 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 1724 | 1746 | POTENTIAL. |

```
FT DISULFID 381 398 BY SIMILARITY.
FT DISULFID 400 409 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 456 467 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 494 505 BY SIMILARITY.
FT DISULFID 499 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 532 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 554 563 BY SIMILARITY.
FT DISULFID 570 580 BY SIMILARITY.
FT DISULFID 575 589 BY SIMILARITY.
FT DISULFID 591 600 BY SIMILARITY.
FT DISULFID 607 618 BY SIMILARITY.
FT DISULFID 612 627 BY SIMILARITY.
FT DISULFID 629 638 BY SIMILARITY.
FT DISULFID 645 655 BY SIMILARITY.
FT DISULFID 650 664 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 682 693 BY SIMILARITY.
FT DISULFID 687 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT DISULFID 720 730 BY SIMILARITY.
FT DISULFID 725 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 855 BY SIMILARITY.
FT DISULFID 857 866 BY SIMILARITY.
FT DISULFID 873 884 BY SIMILARITY.
FT DISULFID 878 893 BY SIMILARITY.
FT DISULFID 895 904 BY SIMILARITY.
FT DISULFID 911 922 BY SIMILARITY.
FT DISULFID 916 931 BY SIMILARITY.
FT DISULFID 933 942 BY SIMILARITY.
FT DISULFID 987 998 BY SIMILARITY.
FT DISULFID 992 1007 BY SIMILARITY.
FT DISULFID 1009 1018 BY SIMILARITY.
FT DISULFID 1025 1036 BY SIMILARITY.
FT DISULFID 1030 1045 BY SIMILARITY.
FT DISULFID 1047 1056 BY SIMILARITY.
FT DISULFID 1063 1074 BY SIMILARITY.
FT DISULFID 1068 1083 BY SIMILARITY.
FT DISULFID 1085 1094 BY SIMILARITY.
FT DISULFID 1101 1122 BY SIMILARITY.
FT DISULFID 1116 1131 BY SIMILARITY.
FT DISULFID 1133 1142 BY SIMILARITY.
FT DISULFID 1149 1160 BY SIMILARITY.
FT DISULFID 1154 1169 BY SIMILARITY.
FT DISULFID 1171 1180 BY SIMILARITY.
FT DISULFID 1187 1198 BY SIMILARITY.
FT DISULFID 1192 1207 BY SIMILARITY.
FT DISULFID 1209 1218 BY SIMILARITY.
FT DISULFID 1225 1244 BY SIMILARITY.
FT DISULFID 1238 1253 BY SIMILARITY.
FT DISULFID 1255 1264 BY SIMILARITY.
FT DISULFID 1271 1284 BY SIMILARITY.
FT DISULFID 1276 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1311 1322 BY SIMILARITY.
FT DISULFID 1316 1334 BY SIMILARITY.
FT DISULFID 1336 1345 BY SIMILARITY.
FT DISULFID 1352 1363 BY SIMILARITY.
FT DISULFID 1357 1372 BY SIMILARITY.
```

```
FT DISULFID 1374 1383 BY SIMILARITY.
FT DISULFID 1391 1403 BY SIMILARITY.
FT DISULFID 1397 1414 BY SIMILARITY.
FT DISULFID 1416 1425 BY SIMILARITY.
FT CARBOHYD 41 41 POTENTIAL.
FT CARBOHYD 888 888 POTENTIAL.
FT CARBOHYD 959 959 POTENTIAL.
FT CARBOHYD 1179 1179 POTENTIAL.
FT CARBOHYD 1241 1241 POTENTIAL.
FT CARBOHYD 1489 1489 POTENTIAL.
FT CARBOHYD 1587 1587 POTENTIAL.
SQ SEQUENCE 2531 AA; 270906 MW; 08B31D2D CRC32;

Query Match 14.9%; Score 920; DB 1; Length 2531;
Best Local Similarity 40.8%; Pred. No. 3,82e+195;
Matches 129; Conservative 68; Mismatches 97; Indels 22; Gaps 17;

Db 718 NECNSNP-CIRGACRDGNGYKCDAPGSGTNCIDINNECESN-PCVNGGTCMDTSG- 774
QY 258 NECVLEPNCIHGTC-N--KPMTCICNBSMGGLXCMODLNTCTNHRPCNKGTCFRTGSL 314
Db 775 YVCTCREGFSGPNCQTINNEC-AS-NPCLNGTCIDD--V-AGYKNCPLPYGTATCEV- 828
QY 315 YTKCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEBHTKYGKHCANMGSGMCEK 374
Db 829 VLAFCATSPCKNSGVCKE--SE-DY--EFSFCVPTGNGGTCEIDINECVKSPCRHGAS 883
QY 375 VLT-CSDPFC-HOGICRNVRPGLSKGGYGCECPYIGYSGPNCIDQLDNCSPNFCINSGS 432
Db 884 CONTNGSYRCACQAGYTRNCESDIDCRPNPCHNGSGCTGVNAAFCDCLPGFGARCE 943
QY 433 CQ-PSG--KCLCPAGFSSTRCETINDDCLGHQCENGGICIMVNDYRCQVPGFHGHS 489
Db 944 EDINECATNPQNGANCCTVDYSTCTCTPFGNHCENNTPDCTESSCFNGTCVDGIN 1003
QY 490 SKVDLCILRCPANGGTCLLNNDYQCTCRAGFTGKDCSVDDIDESSGPHNGGTCMRYN 549
Db 1004 SFTCLCPGFTGSYQ 1019
QY 550 SFEVCANGFRGKOD 565

RESULT 11
ID NTCL_HUMAN STANDARD; PRT; 2444 AA.
AC P46531;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1 PRECURSOR (TRANSLATION-
DE ASSOCIATED NOTCH PROTEIN TAN-1) (FRAGMENT).
GN NOTCH1 OR TANI.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE; 91347367.
RA ELLISEN L.W., BIRD J., WEST D.C., SORENG A.L., REYNOLDS T.C.,
RA SMITH S.D., SKLAR J.;
RL CELL 66:649-661(1991).
CC -1- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN
CC ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION
CC IN SOME T-CELL NEOPLASMS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,
CC BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT
CC IS FOUND MAINLY IN LYMPHOID TISSUES.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 3 COPIES OF A LIN/NOTCH REPEAT.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
DR EMBL; M73980; G338675; -.
DR MIM; 190198; -.

```

| | | | | | | |
|----|---------------------------------------------------------------------|----|----------|------|------|----------------|
| DR | PROSITE; PS00010; ASX_HYDROXYL; 20. | FT | DISULFID | 111 | 127 | BY SIMILARITY. |
| DR | PROSITE; PS00022; EGF_1; 34. | FT | DISULFID | 129 | 138 | BY SIMILARITY. |
| DR | PROSITE; PS01186; EGF_2; 26. | FT | DISULFID | 144 | 135 | BY SIMILARITY. |
| DR | PROSITE; PS01187; EGF_CA; 18. | FT | DISULFID | 149 | 164 | BY SIMILARITY. |
| KW | DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN; | FT | DISULFID | 166 | 175 | BY SIMILARITY. |
| KW | TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN. | FT | DISULFID | 182 | 195 | BY SIMILARITY. |
| FT | SIGNAL | FT | DISULFID | 189 | 204 | BY SIMILARITY. |
| FT | CHAIN | FT | DISULFID | 206 | 215 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 222 | 233 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 227 | 243 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 245 | 254 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 261 | 272 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 266 | 281 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 283 | 292 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 299 | 312 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 306 | 321 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 323 | 332 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 339 | 350 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 344 | 359 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 361 | 370 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 376 | 387 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 381 | 398 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 400 | 409 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 416 | 429 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 423 | 438 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 440 | 449 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 456 | 467 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 461 | 476 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 478 | 487 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 494 | 505 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 499 | 514 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 516 | 525 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 532 | 543 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 537 | 552 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 554 | 563 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 570 | 580 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 575 | 589 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 591 | 600 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 607 | 618 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 612 | 627 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 629 | 638 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 645 | 655 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 650 | 664 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 666 | 675 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 682 | 693 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 687 | 702 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 704 | 713 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 720 | 730 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 725 | 739 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 741 | 750 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 757 | 768 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 762 | 777 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 779 | 788 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 795 | 806 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 800 | 815 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 817 | 826 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 833 | 844 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 838 | 855 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 857 | 867 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 874 | 885 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 879 | 894 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 896 | 905 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 912 | 923 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 917 | 932 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 934 | 943 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 948 | 959 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 988 | 999 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 993 | 1008 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1010 | 1019 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1026 | 1037 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1031 | 1046 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1048 | 1057 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1064 | 1075 | BY SIMILARITY. |
| FT | DOMAIN | FT | DISULFID | 1069 | 1084 | BY SIMILARITY. |


```
FT DOMAIN 191 229 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL). 481 496 FT DISULFID BY SIMILARITY.
FT DOMAIN 231 271 EGF-LIKE 6. 498 507 FT DISULFID BY SIMILARITY.
FT DOMAIN 309 309 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL). 514 525 FT DISULFID BY SIMILARITY.
FT DOMAIN 311 350 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL). 519 534 FT DISULFID BY SIMILARITY.
FT DOMAIN 352 388 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL). 536 545 FT DISULFID BY SIMILARITY.
FT DOMAIN 389 427 EGF-LIKE 10. 552 563 FT DISULFID BY SIMILARITY.
FT DOMAIN 429 470 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL). 557 572 FT DISULFID BY SIMILARITY.
FT DOMAIN 472 508 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL). 574 583 FT DISULFID BY SIMILARITY.
FT DOMAIN 510 546 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL). 590 601 FT DISULFID BY SIMILARITY.
FT DOMAIN 548 584 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL). 612 621 FT DISULFID BY SIMILARITY.
FT DOMAIN 586 622 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL). 626 637 FT DISULFID BY SIMILARITY.
FT DOMAIN 622 656 EGF-LIKE 16. 631 646 FT DISULFID BY SIMILARITY.
FT DOMAIN 658 686 EGF-LIKE 17. 648 655 FT DISULFID BY SIMILARITY.
FT DOMAIN 688 724 EGF-LIKE 18. 662 669 FT DISULFID BY SIMILARITY.
FT DOMAIN 726 762 EGF-LIKE 19. 664 674 FT DISULFID BY SIMILARITY.
FT DOMAIN 764 800 EGF-LIKE 20. 676 685 FT DISULFID BY SIMILARITY.
FT DOMAIN 803 839 EGF-LIKE 21. 692 703 FT DISULFID BY SIMILARITY.
FT DOMAIN 841 877 EGF-LIKE 22. 697 712 FT DISULFID BY SIMILARITY.
FT DOMAIN 878 924 EGF-LIKE 23. 714 723 FT DISULFID BY SIMILARITY.
FT DOMAIN 926 962 EGF-LIKE 24. 730 741 FT DISULFID BY SIMILARITY.
FT DOMAIN 964 1000 EGF-LIKE 25. 741 750 FT DISULFID BY SIMILARITY.
FT DOMAIN 1002 1040 EGF-LIKE 26. 752 761 FT DISULFID BY SIMILARITY.
FT DOMAIN 1042 1081 EGF-LIKE 27. 768 779 FT DISULFID BY SIMILARITY.
FT DOMAIN 1083 1122 EGF-LIKE 28. 773 788 FT DISULFID BY SIMILARITY.
FT DOMAIN 1126 1167 EGF-LIKE 29. 790 799 FT DISULFID BY SIMILARITY.
FT DOMAIN 1168 1282 3 X LIN/NOTCH REPEATS. 799 818 FT DISULFID BY SIMILARITY.
FT REPEAT 1208 1208 LIN/NOTCH 1. 807 818 FT DISULFID BY SIMILARITY.
FT REPEAT 1209 1242 LIN/NOTCH 2. 827 838 FT DISULFID BY SIMILARITY.
FT REPEAT 1243 1282 LIN/NOTCH 3. 829 838 FT DISULFID BY SIMILARITY.
FT DOMAIN 1272 1785 6 X ANK MOTIF REPEATS. 845 856 FT DISULFID BY SIMILARITY.
FT REPEAT 1572 1603 ANK MOTIF 1. 850 865 FT DISULFID BY SIMILARITY.
FT REPEAT 1622 1653 ANK MOTIF 2. 867 876 FT DISULFID BY SIMILARITY.
FT REPEAT 1654 1685 ANK MOTIF 3. 882 903 FT DISULFID BY SIMILARITY.
FT REPEAT 1688 1719 ANK MOTIF 4. 897 912 FT DISULFID BY SIMILARITY.
FT REPEAT 1721 1752 ANK MOTIF 5. 914 923 FT DISULFID BY SIMILARITY.
FT REPEAT 1754 1785 ANK MOTIF 6. 930 941 FT DISULFID BY SIMILARITY.
FT DISULFID 25 38 BY SIMILARITY. 941 950 FT DISULFID BY SIMILARITY.
FT DISULFID 32 48 BY SIMILARITY. 952 961 FT DISULFID BY SIMILARITY.
FT DISULFID 50 59 BY SIMILARITY. 968 979 FT DISULFID BY SIMILARITY.
FT DISULFID 65 77 BY SIMILARITY. 973 988 FT DISULFID BY SIMILARITY.
FT DISULFID 71 100 BY SIMILARITY. 990 999 FT DISULFID BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY. 1006 1019 FT DISULFID BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY. 1011 1028 FT DISULFID BY SIMILARITY.
FT DISULFID 124 140 BY SIMILARITY. 1030 1039 FT DISULFID BY SIMILARITY.
FT DISULFID 142 151 BY SIMILARITY. 1046 1057 FT DISULFID BY SIMILARITY.
FT DISULFID 157 168 BY SIMILARITY. 1051 1069 FT DISULFID BY SIMILARITY.
FT DISULFID 162 177 BY SIMILARITY. 1071 1080 FT DISULFID BY SIMILARITY.
FT DISULFID 179 188 BY SIMILARITY. 1087 1098 FT DISULFID BY SIMILARITY.
FT DISULFID 195 208 BY SIMILARITY. 1092 1110 FT DISULFID BY SIMILARITY.
FT DISULFID 202 217 BY SIMILARITY. 1112 1121 FT DISULFID BY SIMILARITY.
FT DISULFID 219 228 BY SIMILARITY. 1130 1142 FT DISULFID BY SIMILARITY.
FT DISULFID 235 246 BY SIMILARITY. 1136 1155 FT DISULFID BY SIMILARITY.
FT DISULFID 240 259 BY SIMILARITY. 1157 1166 FT DISULFID BY SIMILARITY.
FT DISULFID 261 270 BY SIMILARITY. 1171 711 FT CARBOHYD POTENTIAL.
FT DISULFID 235 246 BY SIMILARITY. 960 960 FT CARBOHYD POTENTIAL.
FT DISULFID 240 259 BY SIMILARITY. 1139 1139 FT CARBOHYD POTENTIAL.
FT DISULFID 261 270 BY SIMILARITY. 1547 1547 FT CARBOHYD POTENTIAL.
FT DISULFID 277 288 BY SIMILARITY. 43 43 FT CONFLICT Q -> R (IN REF. 3).
FT DISULFID 282 297 BY SIMILARITY. 298 298 FT CONFLICT L -> P (IN REF. 3).
FT DISULFID 299 308 BY SIMILARITY. 884 884 FT CONFLICT M -> K (IN REF. 3).
FT DISULFID 315 329 BY SIMILARITY. 927 927 FT CONFLICT M -> V (IN REF. 3).
FT DISULFID 323 338 BY SIMILARITY. 966 966 FT CONFLICT E -> D (IN REF. 3).
FT DISULFID 340 349 BY SIMILARITY. 1245 1245 FT CONFLICT L -> P (IN REF. 3).
FT DISULFID 356 367 BY SIMILARITY. 1489 1489 FT CONFLICT A -> T (IN REF. 3).
FT DISULFID 371 376 BY SIMILARITY. 1549 1549 FT CONFLICT S -> G (IN REF. 3).
FT DISULFID 378 387 BY SIMILARITY. 1688 1688 FT CONFLICT S -> T (IN REF. 3).
FT DISULFID 393 404 BY SIMILARITY. 1836 1838 FT CONFLICT AAA -> APR (IN REF. 3).
FT DISULFID 398 415 BY SIMILARITY. 1964 AA; 206691 MW; 3F16B6B5 CRC32; SQ SEQUENCE
FT DISULFID 417 426 BY SIMILARITY. 12.9%; Score 798; DB 1; Length 1964;
FT DISULFID 433 449 BY SIMILARITY. Best Local Similarity 37.2%; Pred. No. 7.43e-164;
FT DISULFID 443 458 BY SIMILARITY. Matches 123; Conservative 73; Mismatches 110; Indels 25; Gaps 20;
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
```

Mon Jun 15 11:54:32 1998

| | | | | | | | | | |
|--------|----------------------------------------------------------------|-------------------------------------------------------------|---------------|----|--------|------|------|----------------|-----------------------------|
| DB | 270 | GMNPDGCRHOCQNGATIDGL-DR-YTLCPRKMGQCSIEDCEARGPRCNGG | 327 | FT | DOMAIN | 427 | 425 | EGF-LIKE 4. | CALCIUM-BINDING (POTENTIAL) |
| OY | 248 | CQDGMKGAICNEVLEPNCHTGNKFWTICNENMGGLYNODLNTYHN-RP-CKNGG | 305 | FT | DOMAIN | 464 | 500 | EGF-LIKE 6. | |
| DB | 328 | TCQNTA-GSFHCYVSGWAGGACENLDCAAT--CAPSTCID--RV-GSFSCLPFG | 381 | FT | DOMAIN | 501 | 532 | EGF-LIKE 7. | |
| OY | 306 | TCFNTGEGLYTCKACARGYSDDCENELIYSCDADVPQNGGTCTIDEHTHTGYKCHANG | 365 | FT | DOMAIN | 582 | 611 | EGF-LIKE 9. | |
| DB | 382 | RFGGLCHLBDM-CLSPCHVNAQC-STNP-L-T-GSTL-CICOPGYSGTCHODLDECOM | 435 | FT | DOMAIN | 609 | 646 | EGF-LIKE 10. | |
| OY | 366 | WGMKMEERVLNCSDRPCH-QGICRNVRPGLSGKGGYQCECPYSGSPNDLIDLNC-- | 422 | FT | DOMAIN | 687 | 723 | EGF-LIKE 11. | CALCIUM-BINDING (POTENTIAL) |
| DB | 436 | AQGGPSCHEGSGCINTPSSFNCLTAPGYTSRCEADHNKLSOPCHRGSTCDLLATFH | 495 | FT | DOMAIN | 725 | 761 | EGF-LIKE 12. | CALCIUM-BINDING (POTENTIAL) |
| OY | 423 | S---PWPCCINSGGC-0-PSG-KCICPAGSFSGTRETINIDCLGHCQNGGCTDMVNOYR | 476 | FT | DOMAIN | 763 | 800 | EGF-LIKE 13. | CALCIUM-BINDING (POTENTIAL) |
| DB | 496 | CLCPPLGRLCEVEVNECTSNPCLNQACHDLNGFQCLPGLFGARGEKMDDECSST | 555 | FT | DOMAIN | 802 | 838 | EGF-LIKE 15. | CALCIUM-BINDING (POTENTIAL) |
| OY | 477 | CQCVPFHHTHGSKYDCLIRPCANGGTCLNINDYQCCRGAFGTOKDSDYDIDESSG | 536 | FT | DOMAIN | 904 | 940 | EGF-LIKE 16. | CALCIUM-BINDING (POTENTIAL) |
| DB | 556 | PCANGRCRDQPGAFYCECLPGFEGPCEKE | 586 | FT | DOMAIN | 942 | 978 | EGF-LIKE 17. | CALCIUM-BINDING (POTENTIAL) |
| OY | 537 | PCHNGGTGMNRVNSFEVCVANGFRKQCDDE | 567 | FT | DOMAIN | 980 | 1021 | EGF-LIKE 18. | CALCIUM-BINDING (POTENTIAL) |
| RESULT | 15 | STANDARD: | PRT; 2139 AA. | FT | DOMAIN | 1021 | 1243 | EGF-LIKE 19. | |
| ID | CRB-DROME | | | FT | DOMAIN | 1207 | 1243 | EGF-LIKE 20. | |
| AC | P10040: | | | FT | DOMAIN | 1481 | 1517 | EGF-LIKE 21. | |
| DT | 01-MAR-1989 (REL. 10, CREATED) | | | FT | DOMAIN | 1759 | 1795 | EGF-LIKE 22. | |
| DT | 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE) | | | FT | DOMAIN | 1797 | 1833 | EGF-LIKE 23. | CALCIUM-BINDING (POTENTIAL) |
| DT | 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE) | | | FT | DOMAIN | 1835 | 1871 | EGF-LIKE 24. | CALCIUM-BINDING (POTENTIAL) |
| DE | CRUMBS PROTEIN PRECURSOR (95F). | | | FT | DOMAIN | 1915 | 1951 | EGF-LIKE 25. | |
| GN | CRB. | | | FT | DOMAIN | 1953 | 1989 | EGF-LIKE 26. | |
| OS | DROSOPHILA MELANOGASTER (FRUIT FLY). | | | FT | DOMAIN | 1991 | 2029 | EGF-LIKE 27. | CALCIUM-BINDING (POTENTIAL) |
| OC | EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA. | | | FT | DOMAIN | 2030 | 2070 | EGF-LIKE 28. | CALCIUM-BINDING (POTENTIAL) |
| RN | [1] | | | FT | DOMAIN | 271 | 282 | EGF-LIKE 29. | |
| RP | SEQUENCE FROM N.A. | | | FT | DOMAIN | 276 | 291 | BY SIMILARITY. | |
| RC | STRAIN-OREGON-R; TISSUE-EMBRYO; | | | FT | DOMAIN | 293 | 302 | BY SIMILARITY. | |
| RX | MEDLINE; 90263104. | | | FT | DOMAIN | 310 | 321 | BY SIMILARITY. | |
| RL | TEPASS U., THERES C., KNUST E.; | | | FT | DOMAIN | 315 | 331 | BY SIMILARITY. | |
| RL | CELL 61:787-799(1990). | | | FT | DOMAIN | 333 | 342 | BY SIMILARITY. | |
| RN | [2] | | | FT | DOMAIN | 352 | 363 | BY SIMILARITY. | |
| RP | SEQUENCE OF 1663-1955 FROM N.A. | | | FT | DOMAIN | 357 | 374 | BY SIMILARITY. | |
| RX | MEDLINE; 87218537. | | | FT | DOMAIN | 376 | 385 | BY SIMILARITY. | |
| RA | KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D., | | | FT | DOMAIN | 392 | 403 | BY SIMILARITY. | |
| RA | VAESSIN H., CAMPOS-ORTEGA J.A.; | | | FT | DOMAIN | 414 | 424 | BY SIMILARITY. | |
| RL | EMBO J. 6:761-766(1987). | | | FT | DOMAIN | 431 | 451 | BY SIMILARITY. | |
| CC | -I- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA, | | | FT | DOMAIN | 453 | 462 | BY SIMILARITY. | |
| CC | POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL | | | FT | DOMAIN | 468 | 479 | BY SIMILARITY. | |
| CC | POLARITY. IT MAY ACT AS A SIGNAL. | | | FT | DOMAIN | 473 | 488 | BY SIMILARITY. | |
| CC | -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | FT | DOMAIN | 490 | 499 | BY SIMILARITY. | |
| CC | -I- PWM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL). | | | FT | DOMAIN | 505 | 515 | BY SIMILARITY. | |
| CC | -I- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS. | | | FT | DOMAIN | 522 | 531 | BY SIMILARITY. | |
| DR | EMBL; M33753; G552087; ALT_SEQ. | | | FT | DOMAIN | 549 | 562 | BY SIMILARITY. | |
| DR | EMBL; X05144; E1746; -. | | | FT | DOMAIN | 556 | 565 | BY SIMILARITY. | |
| DR | PIR; B26637; B26637. | | | FT | DOMAIN | 571 | 580 | BY SIMILARITY. | |
| DR | PIR; A35672; A35672. | | | FT | DOMAIN | 586 | 597 | BY SIMILARITY. | |
| DR | HSSP; P00743; IAPD. | | | FT | DOMAIN | 591 | 602 | BY SIMILARITY. | |
| DR | FLYBASE: FBgn0000368; crb. | | | FT | DOMAIN | 604 | 610 | BY SIMILARITY. | |
| DR | PROSITE; PS00010; ASX_HYDROXYL; 15. | | | FT | DOMAIN | 613 | 624 | BY SIMILARITY. | |
| DR | PROSITE; PS00022; EGF_1; 26. | | | FT | DOMAIN | 618 | 634 | BY SIMILARITY. | |
| DR | PROSITE; PS01186; EGF_2; 17. | | | FT | DOMAIN | 636 | 645 | BY SIMILARITY. | |
| DR | PROSITE; PS01187; EGF_CA; 15. | | | FT | DOMAIN | 652 | 664 | BY SIMILARITY. | |
| KW | DIFFERENTIATION; REPEAT; EGF-LIKE DOMAIN; TRANSMEMBRANE; | | | FT | DOMAIN | 659 | 673 | BY SIMILARITY. | |
| KW | GLYCOPROTEIN; SIGNAL; PHOSPHORYLATION. | | | FT | DOMAIN | 675 | 684 | BY SIMILARITY. | |
| FT | SIGNAL | 1 | 90 | FT | DOMAIN | 691 | 702 | BY SIMILARITY. | |
| FT | CHAIN | 91 | 2139 | FT | DOMAIN | 767 | 778 | BY SIMILARITY. | |
| FT | DOMAIN | 91 | 2084 | FT | DOMAIN | 772 | 787 | BY SIMILARITY. | |
| FT | TRANSMEM | 2085 | 2111 | FT | DOMAIN | 789 | 799 | BY SIMILARITY. | |
| FT | DOMAIN | 2112 | 2139 | FT | DOMAIN | 806 | 817 | BY SIMILARITY. | |
| FT | DOMAIN | 267 | 303 | FT | DOMAIN | 811 | 826 | BY SIMILARITY. | |
| FT | DOMAIN | 306 | 343 | FT | DOMAIN | 828 | 837 | BY SIMILARITY. | |
| FT | DOMAIN | 348 | 386 | FT | DOMAIN | 844 | 855 | BY SIMILARITY. | |
| FT | DOMAIN | | | FT | DOMAIN | 849 | 890 | BY SIMILARITY. | |

QY 493 DCLIRPCANGCTCLNNDYQCTCRAGFTGKDCSDVIDECSSGP 537

RESULT 3
ID P87357; PRELIMINARY; PRT; 717 AA.
AC P87357;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DELTAD TRANSMEMBRANE PROTEIN PRECURSOR.
GN DELTAD.
OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTRICHITES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RA DORNSEITER P., TAKKE C., CAMPOS-ORTEGA J.A.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Y11760; E307461; -
DR PROSITE: PS01187; EGF_CA; 2.
KW SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 717 DELTAD TRANSMEMBRANE PROTEIN.
SQ SEQUENCE 717 AA; 79061 MW; 5CC32ECA CRC32;

Query Match 30.8%; Score 1897; DB 12; Length 717;
Best Local Similarity 47.7%; Pred. No. 0.0e+00;
Matches 251; Conservative 96; Mismatches 157; Indels 22; Gaps 18;

DB 5 MINVILC-VNISQGFSGVPELKLQEFVNKKYVGNANCKG-SAAEGHOC-BECKTEFR 60
QY 7 LTLAFTFYIVYVHSSGSEFLKLFYNSDHRGNEGRCCSGSDGATGKLGSCKTR 66
DB 61 ICLKHYQANVSPDPCTYGSVAIVPVLSNSFVPESS-FPPSSFTNPPIPAFGFTWPTF 118
QY 67 VCLKHQATIDTTSQCTYGVDPILIGENSVNLDAQRFQKFTNPIDPFSSWPTF 126
DB 119 SLTIEALHTDSTDLSTENPDRLISKMTORHLTVGEEMSDLOVGGTELYSTRFYCD 178
QY 127 SLIVEAMH-DTNNSGNARTKLLIQRLLVQVLEVSSEMKTKSEQYTSLEYDFRYCD 185
DB 179 EHYEGGSCVFCRPDRDAFGHFTGSEKGEVNCNPKMGKGPCTEPLCPGDEHGFCDK 238
QY 186 LNTYSGCAFCRPRDRDSFGHSTCEGTEILCTMGQDCHIRPKAKGC-E-HGHCCKP 243
DB 239 GECKCRVGFSGKYDDCIRYPGCLHGTQCPMOCNCOEGMGLFCNODLNTCTHHRPC 298
QY 244 NQVCQGLGMKALCNCEVLEPNCIHGTCKNPKWICICNEGGLYCNDLNTCTHHRPC 303
DB 299 GATCTNTGGSYTCSPRGFTGDSCEIEVNEC-SG-SPCRNGSCD-LF-NT-YSCTCP 353
QY 304 GGCFTNGEGLYTCCKAPYSGSDCENETYSADADVNPQNGGTCIDEPHTKTYGKCA 363
DB 354 PGFTYGRNCELSAMTCAAPCENGG-HC-AD-NPEG-GYFCQCPMGTAAGFCEKKIDHC 408
QY 364 NGMSGKCEKRYLTCSPKHOGICRNVRPGISRGOGYQCECPIDYSGPCMDQLDNC 423
DB 409 SNPCSDAOCGLDLYDYLCCPPEFTGHCEHNDIDECATYCCONGGCTODLSYTCPC 468
QY 424 PNCINGSC-QPSGK--CICPAGFSGTRETNIIDCLGHCEGNGGCTIDVNOIRCOV 480
DB 469 PGYTKNCTSAVNKCLNPNCHNGATCEHMDRNYCACIPGYGANG 514
QY 481 PGRHGTGSSKVDCLIRPCANGCTCLNNDYQCTCRAGFTGKDC 526

RESULT 4
ID 000548; PRELIMINARY; PRT; 723 AA.
AC 000548;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)

DE DELTA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MANN R.S., GRAY G.E., HENRIQUE D., ISH-HOROWICZ D.,
RA ARAVANIS-TSAKONAS S.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF003522; G2197069; -
DR PROSITE: PS01187; EGF_CA; 1.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 723 AA; 77956 MW; A1D48BDB CRC32;

Query Match 30.3%; Score 1865; DB 2; Length 723;
Best Local Similarity 48.4%; Pred. No. 0.0e+00;
Matches 261; Conservative 86; Mismatches 171; Indels 21; Gaps 17;

DB 4 RCALAVIALSCVMSGVPELKLQEFVNKKGLGNRCCRGGA-GPP-C--ACRTF 59
QY 5 KCLTAFICFTYIVYVHSSGSEFLKLFYNSDHRGNEGRCCSGSDGATGKLGSCKTR 64
DB 60 FRVCLKHQASVSPDPCTYGSVAIVPVLSNSFVPESS-ADSAFNPRIREFFTWPG 118
QY 65 FRVCLKHQATIDTTSQCTYGVDPILIGENSVNLDAQRFQKFTNPIDPFSSWPTF 124
DB 119 TFSLLIEALHTDSTDLSTENPDRLISKMTORHLTVGEEMSDLOVGGTELYSTRFY 178
QY 125 TFSLLIEAMH-DTNNSGNARTKLLIQRLLVQVLEVSSEMKTKSEQYTSLEYDFRYT 183
DB 179 CDEHYEGGSCVFCRPDRDAFGHFTGSEKGEVNCNPKMGKGPCTEPLCPGDEHGFCD 238
QY 184 CDLNTYSGCAFCRPRDRDSFGHSTCEGTEILCTMGQDCHIRPKAKGC-E-HGHC 241
DB 239 KPECKCRVGFSGKYDDCIRYPGCLHGTQCPMOCNCOEGMGLFCNODLNTCTHHRPC 298
QY 242 KPNQVCQGLGMKALCNCEVLEPNCIHGTCKNPKWICICNEGGLYCNDLNTCTHHRPC 301
DB 299 KNGATCTNTGGSYTCSPRGFTGATGELGIDCDP--SPCRNGSCD-LF-NS-YSCCT 353
QY 302 KNGGCFNTGEGLYTCCKAPYSGSDCENETYSADADVNPQNGGTCIDEPHTKTYGKCH 361
DB 354 CPGFTYGRNCELSAMTCAAPCENGG-R-C--SD-SPDG-GYSCRCRPGVSGNCEKKIDY 408
QY 362 CANGSGMCEKRYLTCSDKPCPHOGICRNVRPGISRGOGYQCECPIDYSGPCMDQLDN 421
DB 409 CSSPSCSNAGCAVDLGDYVLCRCQAGFSGRHCDVNDVDCASSPANGGTCRQGVNDFCT 468
QY 422 CSPNCSINGSCQPSGK---CICPAGFSGTRETNIIDCLGHCEGNGGCTIDVNOIRCO 478
DB 469 CPGFTYGRNCELSAMTCAAPCENGGTCRQGVNDFCT 527
QY 479 CVPGRHGTGSSKVDCLIRPCANGCTCLNNDYQCTCRAGFTGKDCSDVIDECSSGP 537

RESULT 5
ID 091902; PRELIMINARY; PRT; 721 AA.
AC 091902;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE X-DELTA-1.
OS XENOPUS LAEVIS (AFRICAN CLAMPED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE; 95319507.
RA HENRIQUE D., ADAM J., MYAT A., CHITNIS A., LEWIS J., ISH-HOROWICZ D.,
RL NATURE 375:787-790(1995).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE; 95319503.
RA CHITNIS A.B., HENRIQUE D., LEWIS J., ISH-HOROWICZ D., KINTNER C.R.;


```

QY 7 LTAFAICTVIVYH-SSGSFELRLKYPFNDHGRNDEGRCCSGESDAGTKCL-GSKCTR 64
DB 75 FVCLKEQVATIDTTSOCTYGDVITPILGENSEVNLDAQRFONKGFTHPIOPFESFWPG 130
QY 65 FVCLKEQVATIDTTSOCTYGDVITPILGENSEVNLDAQRFONKGFTHPIOPFESFWPG 124
DB 131 SYTLLEAM-DSSND-TVOPDSI-IEKASHSGMINPSROMOTLKONTGAHEYOIRVTC 187
QY 125 TFSLIVEAMHDNTNSGNARTNKLILQRLVQOVLEVSSEMTKNSQYTSLEYDRVTC 184
DB 188 DYYVGFGENKCRPRDDFFGHYACDQNGNKTCEMWMGPECNRAICRGCSPKHSGCTL 247
QY 185 DLNYYGSGCAKFCRPRDDFFGHYACDQNGNKTCEMWMGPECNRAICRGCSPKHSGCTL 242
DB 248 PGDCRCOYGMOGILYCDKCPHPCGVHIGICNEPWCCLCTNMGGQLCDKDLNCTGTHQPC 307
QY 243 PNOVCQQLGWMKALCNECILEPNCIHGTCKNPMWTCICNMGWGLCNDLNTCTHNRPK 302
DB 308 NGGTCMTGPDYKXGSCGEGYSGPNCIEAHACLSD--PCHNRGSC-KE--TSLGFECEC 362
QY 303 NGGTCMTGPDYKXGSCGEGYSGPNCIEAHACLSD--PCHNRGSC-KE--TSLGFECEC 362
DB 363 SPGWTGPTCSTINDCSPNHCSSHGTCOD---L-VNG--FKVCPPONTGTCOLDANE 415
QY 363 ANGWSGKMCERKVLTCSDKPC-HOGICRNVPRGLSGKGGYQCEPPIGYSGPCDLOLDN 421
DB 416 CEAKPCVNAKSCKNLIASYCDCLPGWMGONCININDCLG-OCONDAACRDLVNGYRCI 474
QY 422 CSPNRCINGSCQP-SGK--CICPAGFSGTRETINIDCLGHCENGGTCLDMVQYRCQ 478
DB 475 CPBGYAGDHCEBIDECAANPCLNGHCONEINRFQCLPTGSGNLCOLDIDYCEBPNPC 534
QY 479 CVPGFHGTCSKXVDCLIRPCANGGTCLNINDYQCTCRAGFTGDKSVDIDECSSGPC 538
DB 535 ONGAOCYNRASDYFCCKPEDEYCKNC 560
QY 539 HNGGTCTMNRVNSFEVCANGFERGKOC 564

```

```

RESULT 8
ID 014902 PRELIMINARY; PRT; 1218 AA.
AC 014902;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE PROTEIN JAGGED 1.
GN JAG1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95211842.
RA LINDELL C.E., SHAMBER C.J., BOULTER J., WEINMASTER G.;
RL CELL 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA BASH J., ZONG W.-X., GELINAS C.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF028593; G2599082; -
DR PROSITE: PS01187; EGF_CA: 8.
KM TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 1218 AA; 133798 MW; 07B97EE3 CRC32;

```

Query Match 27.4%; Score 1691; DB 2; Length 1218;
 Best Local Similarity 41.7%; Pred. No. 0.00e+00;
 Matches 236; Conservative 109; Mismatches 192; Indels 29; Gaps 19;

```

DB 16 LTAALIC-ALRAKVCASGOFELEILSMQVNGELONGCGGARNPDRCRTRECDTY 74
QY 7 LTAFAICTVIVYH-SSGSFELRLKYPFNDHGRNDEGRCCSGESDAGTKCL-GSKCTR 64
DB 75 FVCLKEQVATIDTTSOCTYGDVITPILGENSEVNLDAQRFONKGFTHPIOPFESFWPG 130

```

```

QY 65 FVCLKEQVATIDTTSOCTYGDVITPILGENSEVNLDAQRFONKGFTHPIOPFESFWPG 124
DB 131 SYTLLEAM-DSSND-TVOPDSI-IEKASHSGMINPSROMOTLKONTGAHEYOIRVTC 187
QY 125 TFSLIVEAMHDNTNSGNARTNKLILQRLVQOVLEVSSEMTKNSQYTSLEYDRVTC 184
DB 188 DYYVGFGENKCRPRDDFFGHYACDQNGNKTCEMWMGPECNRAICRGCSPKHSGCTL 247
QY 185 DLNYYGSGCAKFCRPRDDFFGHYACDQNGNKTCEMWMGPECNRAICRGCSPKHSGCTL 242
DB 248 PGDCRCOYGMOGILYCDKCPHPCGVHIGICNEPWCCLCTNMGGQLCDKDLNCTGTHQPC 307
QY 243 PNOVCQQLGWMKALCNECILEPNCIHGTCKNPMWTCICNMGWGLCNDLNTCTHNRPK 302
DB 308 NGGTCMTGPDYKXGSCGEGYSGPNCIEAHACLSD--PCHNRGSC-KE--TSLGFECEC 362
QY 303 NGGTCMTGPDYKXGSCGEGYSGPNCIEAHACLSD--PCHNRGSC-KE--TSLGFECEC 362
DB 363 SPGWTGPTCSTINDCSPNHCSSHGTCOD---L-VNG--FKVCPPONTGTCOLDANE 415
QY 363 ANGWSGKMCERKVLTCSDKPC-HOGICRNVPRGLSGKGGYQCEPPIGYSGPCDLOLDN 421
DB 416 CEAKPCVNAKSCKNLIASYCDCLPGWMGONCININDCLG-OCONDAACRDLVNGYRCI 474
QY 422 CSPNRCINGSCQP-SGK--CICPAGFSGTRETINIDCLGHCENGGTCLDMVQYRCQ 478
DB 475 CPBGYAGDHCEBIDECAANPCLNGHCONEINRFQCLPTGSGNLCOLDIDYCEBPNPC 534
QY 479 CVPGFHGTCSKXVDCLIRPCANGGTCLNINDYQCTCRAGFTGDKSVDIDECSSGPC 538
DB 535 ONGAOCYNRASDYFCCKPEDEYCKNC 560
QY 539 HNGGTCTMNRVNSFEVCANGFERGKOC 564

```

```

RESULT 9
ID 015122 PRELIMINARY; PRT; 1218 AA.
AC 015122;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED1.
GN JAG1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA ODA T., ELKAHLOUN A.G., PIKE B.L., OKAJIMA K., KRANTZ I.D., GENIN A.,
RA PICCOLI D.A., MELTZER P.S., SPINNER N.B., COLLINS F.S.,
RA CHANDRASEKHARAPPA S.C.;
RL NAT. GENET. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA ODA T., ELKAHLOUN A.G., MELTZER P.S., CHANDRASEKHARAPPA S.C.;
RL GENOMICS 43:376-379(1997).
DR EMBL: AF003837; G2228793; -
DR PROSITE: PS01187; EGF_CA: 8.
KM GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 1218 AA; 133858 MW; 20F471DB CRC32;

```

Query Match 27.4%; Score 1687; DB 2; Length 1218;
 Best Local Similarity 41.5%; Pred. No. 0.00e+00;
 Matches 235; Conservative 110; Mismatches 192; Indels 29; Gaps 19;

```

DB 16 LTAALIC-ALRAKVCASGOFELEILSMQVNGELONGCGGARNPDRCRTRECDTY 74
QY 7 LTAFAICTVIVYH-SSGSFELRLKYPFNDHGRNDEGRCCSGESDAGTKCL-GSKCTR 64
DB 75 FVCLKEQVATIDTTSOCTYGDVITPILGENSEVNLDAQRFONKGFTHPIOPFESFWPG 124

```

| | | | |
|-----|----|----------------------------------------------------------------|-----|
| 185 | Qy | DLNYYGSGCAKFRPRDRDSTGHSHTCSETGHEIICLTGWQGDYCHIPKCAKCE--HGHCDK | 242 |
| 248 | Db | PGDCRCQYWGQGLYCDRCRIPHPGCVHGHTCNEPQCLCETNMGQQLCDKOLNYCTGHQPC | 307 |
| 243 | Qy | PNQVCVLGWKKGALCNECVLEPNCIIHGTCNKPTWCICNCGWGLYCNODLNYCTNRHPCK | 302 |
| 308 | Db | NRGTCSTWGPDPKQYQCSPEGYSGPNCBIAHACLSD--PCHNRGSC-KE--TSSGFEECE | 362 |
| 303 | Qy | NGGTCFTWTEGLYTKCAPGYSGDGCSENIYSCDADVNPQCGGTCTIDEPTHYTGKYCHC | 362 |
| 363 | Db | SPGWTGTCTSTNDDCSPPNCSHGGTCQD---L-VNG--FKCVCPQWTKTKCOLDANE | 415 |
| 363 | Qy | ANGWSGKMEKVLTCISDKPC-HQGICRNVRPGLSGSGQGYQCECPIGYSGPMCDLQD | 421 |
| 416 | Db | CEAKPCVNAWSCNKLIAYYCDCLPGWMQGNCDININDCLG-QCQNDASCRDLVNGYRCI | 474 |
| 422 | Qy | CSPNFCINGSCQP-SGK--CICPAGEFSGYFCTETNDDCLGHQCEGGGTCTIDMVNYRCQ | 478 |
| 475 | Db | CPPGYAGDHCDERDIDECASNPLNGGHGHCONEINRFQCLCPTGFSGNLCQLDIDYCEPNPC | 534 |
| 479 | Qy | CVPGFHTGHCSSVDKLCILRPCANGGTCLNLNNDYQCTCRAGFTGKDCSVDIDECSSGPC | 538 |
| 535 | Db | QNGAQCYNRASTDYCKCPEDYEGKNC | 560 |
| 539 | Qy | HNGTCMNRVNSFFECVANGFRGQC | 564 |

| | | |
|--------|----------------------------------------------------------------|----------------------------|
| RESULT | 11 | PRELIMINARY; PRT; 1227 AA. |
| AC | P78504 | |
| ID | P78504; | |
| DT | 01-MAY-1997 (TREMBLREL. 03, CREATED) | |
| DT | 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE) | |
| DT | 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE) | |
| DE | JAGGED 1 (TRANSMEMBRANE PROTEIN JAGGED). | |
| GN | HJ1. | |
| OS | HOMO SAPIENS (HUMAN), | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; | |
| OC | EUTHERIA; PRIMATES. | |
| CC | [1] | |
| RN | SEQUENCE FROM N.A. | |
| RP | MEDLINE: 95211842. | |
| EX | LINDSELL C.E., SHAWDER C.J., BOULTER J., WEINMASTER G.; | |
| RA | CELL 80:909-917(1995). | |
| RL | [2] | |
| RN | SEQUENCE FROM N.A. | |
| RP | LI L., DENG Y., BANTA A.B., HOOD L.; | |
| RA | SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS. | |
| RL | [3] | |
| RN | SEQUENCE OF 14-1227 FROM N.A. | |
| RP | MEDLINE: 97115768. | |
| EX | ZIRIN A.B., PEPPER M.S., MCMAHON G., NGUYEN F., MONTESANO R., | |
| RA | MACIAG T.; | |
| RA | J. BIOL. CHEM. 271:32499-32502(1996). | |
| RL | [4] | |
| RN | REVIEWS TO 14-1227. | |
| RP | | |

```

RA ZIMRIN A.B., NGUYEN F., MACIANG T.;;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U73936; G1695274; -.
DR EMBL; U77720; E319096; -.
KW TRANSMEMBRANE.
FT CONFLICT 1187 1227 QRHADTKPLDKOTGOORLGKCPKPNGVHRIADRGHCRR
FT -> NGPFTKHPNWTNKQDRDLDSAQSLNREIYV
FT (IN REF. 1 AND 2).
SQ SEQUENCE 1227 AA; 134770 MW; 5D300B81 CRC32;

Query Match 27.4%; Score 1691; DB 2; Length 1227;
Best Local Similarity 41.7%; Pred. No. 0.00e+00;
Matches 236; Conservative 109; Mismatches 192; Indels 29; Gaps 19;

Db 16 LLLALLC-ALRAKVCGASGFELILSMQNVNVELQNGCCGGARNPDGDKRCTRDCSDIY 74
QY 7 LITAFICFTVIVQVH-SGGSPELRLKLFYSNDHGDRDNRCCSGESDGATKCGI GSKCTR 64

```


Db 181 rvtcdlnyggscakfcrprddsfghstcsetgeiicltgwgdychipk 230
QY 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPK 230

RESULT 2

ID R28960 standard; Protein; 833 AA.
AC R28960,1993 (first entry)
DT 01-APR-1993
DE Delta D11.
KW Human; Notch; plasmid; cDNA; clone; D11; expression library; PCR;
KW polymerase chain reaction; primer; cloning vector; Delta; Serrate;
KW neurogenic; topothymic; homotypic; heterotypic; differentiation;
KW quantitation; antibody.
OS Homo sapiens.
PN W09219734-A.
PD 12-NOV-1992.
PF 01-MAY-1992; U03651.
PR 03-MAY-1991; US-695189.
PR 14-NOV-1991; US-791923.
PA (INDV) UNIV INDIANA FOUND.
PA (OYIA) UNIV YALE.
PI Artavanis-Tsakonas S, Blaumueller CM, Fehon RG, Muskavitch MAT;
PI Rebay I, Shepard SB;
DR WPI: 92-398861/48.
DR N-PSDB: Q30997.
PT Human Notch and Delta DNA and protein sequences - used for study
PT and manipulation of differentiation processes
PS Claim 50; Fig 13; 239pp; English.
CC The sequence given is encoded by the nucleotide sequence of human
CC Delta gene contained in plasmid cDNA clone D11. A human expression
CC library was constructed and screening assays were carried out on to
CC select for the expressed Delta product. Alternatively the sequences
CC could be isolated by amplification using polymerase chain reaction
CC (PCR) primers. The isolated gene may be inserted into a cloning
CC vector and expressed. The Delta gene and also the Notch and Serrate
CC neurogenic genes are designated "topothymic" genes. The proteins
CC they encode are involved in specific homo- or heterotypic interactions
CC crucial to differentiation. The quantitation of mRNA for human Notch
CC and Delta and adhesive molecules, and study of its expression are
CC possible using the DNA and antibodies raised against the Notch and
CC Delta proteins.
SQ Sequence 833 AA;

Query Match 100.0%; Score 1734; DB 6; Length 833;
Best Local Similarity 100.0%; Pred. No. 2.83e-176;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mbwtkclltaficftvqvghsgsfelrklyfndhgrdnegrcsgsdgatkclgs 60
QY 1 MHWIKCLLTAFICFTVIVQVHSGSFEELRKLYFNDHGRDNegrcsgsdgatkclgs 60
Db 61 cktfrvclkhqyatidtsqctygdvtpilgensvnltdaqrfnkgfntpqpfsf 120
QY 61 CKTRFVCLKHQYATIDTSQCTYGDVTPILGENSVNLTDQRFNKGFTNPQPFSE 120
Db 121 swpgtflsiveahwdnnsgnartnklllqrlvqvlevssekwnksesqytsleydf 180
QY 121 SWPGTFLSIVEAHDNNSGNARTNKLRLVQVLEVSESWKNKSESYTSLEYDF 180
Db 181 rvtcdlnyggscakfcrprddsfghstcsetgeiicltgwgdychipk 230
QY 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPK 230

RESULT 3

ID W11719 standard; Protein; 727 AA.
AC W11719;
DT 28-APR-1997 (first entry)
DE C-Delta-1 polypeptide.
KW C-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;

KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy.
OS Gallus sp.
FH Key Location/Qualifiers
FT domain 184..228
FT /label= DSL
FT domain 229..261
FT /label= EGF1
FT domain 262..292
FT /label= EGF2
FT domain 293..332
FT /label= EGF3
FT domain 333..370
FT /label= EGF4
FT domain 371..409
FT /label= EGF5
FT domain 410..447
FT /label= EGF6
FT domain 448..485
FT /label= EGF7
FT domain 486..523
FT /label= EGF8
FT domain 524..534
FT /label= EGF9
FT domain 555..579
FT /label= TM
FT /note= "transmembrane domain"

PN W09701571-A1.

PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PA (OYIA) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
PI Lewis J;
DR WPI: 97-100159/09.
DR N-PSDB: T58897.
PT New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PS Disclosure; Fig 2; 135pp; English.
CC C-Delta-1 polypeptide (W11719) is the chick homologue of Drosophila
CC Delta, a protein that binds to Notch protein. Expression of
CC C-Delta-1 correlates with onset of neurogenesis. The C-delta-1
CC amino acid sequence was deduced from a cDNA clone (T58897) obtd.
CC from chick stage 4-6 embryos. An alternatively spliced variant
CC (W00876) was also isolated, and mouse (W11720) and human (W11721-
CC 38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC can be used to treat or prevent disorders characterised by
CC increased Notch activity, such as cervical, breast, lung or colon
CC cancer, melanoma or seminoma, and nervous system disorders or to
CC promote tissue regeneration and repair.
SQ Sequence 727 AA;

Query Match 44.7%; Score 775; DB 21; Length 727;
Best Local Similarity 49.1%; Pred. No. 2.25e-70;
Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1;

Db 20 qvdsgvfelkqefvknkgllsnrncrgggpggagqgqcdctfrrvclkhqyasvp 79
QY 19 QVHSSGSFELRLKYFNDHGRDNegrcsgsdgatkclgscktrfrrvclkhqyatid 78
Db 80 epctvgysaitpvgansfsvpdagadpafsnpirfpgftwptfslilcalhtdsp 139
QY 79 TSQCTYGDVTPILGENSVNLTDQRFNKGFTNPQPFSESWPGTFLSIVEAHW-DTN 137
Db 140 ddttenperlisrlatqrhlavgeewsdhssgrtdlkysvrfvcdelyegcgvfc 199
QY 138 NSGNARTNKLILQRLVQVLEVSESWKNKSESYTSLEYDFRVTCDLNYGSGCAKFC 197
Db 200 rprddrfghftcgergekvncnpgwkqyctep 231
QY 198 RPRDDSFHSTCSETGEIICLTGWOGDYCHIP 229

| | | |
|---------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| ID | 4 | RESULT |
| WT00876 | standard; Protein; 740 AA. | |
| AC | M00876: | |
| DT | 28-APR-1997 (first entry) | |
| DE | C-delta-1 polypeptide (alternatively spliced variant). | |
| KW | C-beta-1; cell proliferation; nervous system disorder; | |
| KM | tissue regeneration; Notch; cervix cancer; breast cancer; | |
| KV | lung cancer; colon cancer; melanoma; seminoma; | |
| KW | neurogenesis; therapy. | |
| OS | Gallus sp. | |
| FH | Key | Location/Qualifiers |
| FT | domain | 184..228 |
| FT | /label= DSL | |
| FT | domain | 229..261 |
| FT | /label= EGF1 | |
| FT | domain | 262..292 |
| FT | /label= EGF2 | |
| FT | domain | 293..332 |
| FT | /label= EGF3 | |
| FT | domain | 333..370 |
| FT | /label= EGF4 | |
| FT | domain | 371..409 |
| FT | /label= EGF5 | |
| FT | domain | 410..447 |
| FT | /label= EGF6 | |
| FT | domain | 448..485 |
| FT | /label= EGF7 | |
| FT | domain | 486..523 |
| FT | /label= EGF8 | |
| FT | domain | 524..534 |
| FT | /label= EGF9 | |
| FT | domain | 555..579 |
| FT | /label= TM | |
| ET | /note= "transmembrane domain" | |
| PN | WO9701571-A1. | |
| PD | 16-JAN-1997. | |
| PF | 28-JUN-1996: U11178. | |
| PR | 28-JUN-1995: US-000589. | |
| PA | (IMCR) IMPERIAL CANCER RES TECHNOLOGY. | |
| PA | (UYVA) UNIV YALE. | |
| PI | Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D; | |
| PI | Lewis J; | |
| DR | WPt: 97-100159/09. | |
| DR | N-PsDB: T58898. | |
| PT | New vertebrate Delta protein, DNA and antibodies - for treating and | |
| PT | preventing cancer, nervous system disorders and for tissue | |
| PT | regeneration | |
| PS | Disclosure: Fig 2: 135pp: English. | |
| CC | C-delta-1 polypeptide (W00876) is the chick homologue of Drosophila | |
| CC | Delta, a protein that binds to Notch protein. Expression of | |
| CC | C-Delta-1 correlates with onset of neurogenesis. The C-delta-1 | |
| CC | amino acid sequence was deduced from a cDNA clone (Y58898) obtd. | |
| CC | from chick stage 4-6 embryos. A shorter version (W58877) of | |
| CC | C-delta-1, lacking the 12 C-terminal amino acids of the longer | |
| CC | version, was also isolated, and mouse (W11720) and human (W11721- | |
| CC | 38) Delta-1 polypeptides have been identified. Delta-1 proteins | |
| CC | can be used to treat or prevent disorders characterised by | |
| CC | increased Notch activity, such as cervical, breast, lung or colon | |
| CC | cancer, melanoma or seminoma, and nervous system disorders or to | |
| CC | promote tissue regeneration and repair. | |
| SO | Sequence 740 AA: | |
| Query Match | 44.7%: Score 775; DB 21; Length 740; | |
| Best Local Similarity | 49.1%: Pred.No.2,25e-70; | |
| Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1 | | |
| Db | 20 gvdsgvfelekqefvnkkyllsnrncrrgpgypgaqqgccktfifvcikhygasvsp 79 : : : : : :: 19 GVHSSGSEFLRLKRYNSNDHGRIENEGRCSCSGSDGATGCLCSCKTRFVLCMKHYQARTIDT 78 | |
| Db | 80 eppectygsaicrpvygansfsvpdggagdaafpsmpirpfvgftfpgtcsliaelhtcdsp 139 | |

[illegible]

Matches 87; Conservative 50; Mismatches 76; Indels 10; Gaps 10;

```
Db 16 lllallic-alrakvcgagqgflellsmqnvngelqngccgarnpgdrktrdeccty 74
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 7 LTTAFICFTVIVQVH-SGSGFELRLKYPFSDHGRDNEGRCCSGESDGTGKCL-GSCKTR 64
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

- 75 fkvclkeysvrtagpcsfsgstpvigntgnlfnkas-r-gn-d-prvlpfswfawpr 130
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 65 FRVCLKHQATIDTTSQCTYGDVITPILGNSVNLTDQRFQNGKFTNPQPFPSFSPG 124
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

Db 131 sytlilveav-dssnd-tvpqdsi-iekashsgminpsrqwtlqntgvahfeyqirvtc 187
  :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 125 TFSLIVAWHDTNNSGNARTKLLIQRLLVQVLEVSSEWTKNSQYTSLEYDFRVTG 184
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 188 ddyvyygfcnkfcrprddffghyacdngnktcmegwmpcnc 230
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 185 DLNYGSGCAFCRPRDPSFGHSTCSETGEIICLTGWGODYCH 227
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

RESULT 11

```
ID W18354 standard; protein; 1218 AA.
AC W18354;
DE 11-FEB-1998 (first entry)
KW Proliferation and differentiation suppression polypeptide.
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein /label= Signal
FT 32..1218 /label= Differentiation_suppression_protein
FT W09719172-A1.
PD 29-MAY-1997..
PF 15-NOV-1996; J03356.
PR 30-NOV-1995; JP-311811.
PR 17-NOV-1995; JP-299611.
PA (ASAH) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR N-PSDB; T70175.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 15; Page 83-91; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1218 AA;
```

Query Match 33.4%; Score 579; DB 25; Length 1218;

Best Local Similarity 38.1%; Pred. No. 3.94e-49;

Matches 85; Conservative 52; Mismatches 76; Indels 10; Gaps 8;

```
Db 16 lllallic-alrakvcgagqgflellsmqnvngelqngccgarnpgdrktrdeccty 74
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 7 LTTAFICFTVIVQVH-SGSGFELRLKYPFSDHGRDNEGRCCSGESDGTGKCL-GSCKTR 64
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

- 75 fkvclkeysvrtagpcsfsgstpvigntgnlfnkas-r-gn-d-prvlpfswfawpr 130
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 65 FRVCLKHQATIDTTSQCTYGDVITPILGNSVNLTDQRFQNGKFTNPQPFPSFSPG 124
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

Db 131 sytlilveav-dssnd-tvpqdsi-iekashsgminpsrqwtlqntgvahfeyqirvtc 187
  :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 125 TFSLIVAWHDTNNSGNARTKLLIQRLLVQVLEVSSEWTKNSQYTSLEYDFRVTG 184
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 188 ddyvyygfcnkfcrprddffghyacdngnktcmegwmpcnc 230
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 185 DLNYGSGCAFCRPRDPSFGHSTCSETGEIICLTGWGODYCH 227
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

Query Match 33.4%; Score 579; DB 25; Length 1218;

Best Local Similarity 38.1%; Pred. No. 3.94e-49;

Matches 85; Conservative 52; Mismatches 76; Indels 10; Gaps 8;

```
Db 16 lllallic-alrakvcgagqgflellsmqnvngelqngccgarnpgdrktrdeccty 74
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 7 LTTAFICFTVIVQVH-SGSGFELRLKYPFSDHGRDNEGRCCSGESDGTGKCL-GSCKTR 64
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

- 75 fkvclkeysvrtagpcsfsgstpvigntgnlfnkas-r-gn-d-prvlpfswfawpr 130
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 65 FRVCLKHQATIDTTSQCTYGDVITPILGNSVNLTDQRFQNGKFTNPQPFPSFSPG 124
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

Db 131 sytlilveav-dssnd-tvpqdsi-iekashsgminpsrqwtlqntgvahfeyqirvtc 187
  :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 125 TFSLIVAWHDTNNSGNARTKLLIQRLLVQVLEVSSEWTKNSQYTSLEYDFRVTG 184
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 188 ddyvyygfcnkfcrprddffghyacdngnktcmegwmpcnc 230
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 185 DLNYGSGCAFCRPRDPSFGHSTCSETGEIICLTGWGODYCH 227
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

RESULT 12

```
ID W18351 standard; protein; 1036 AA.
AC W18351;
DE 11-FEB-1998 (first entry)
KW Proliferation and differentiation suppression polypeptide.
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein /label= Signal
FT 32..1036 /label= Differentiation_suppression_protein
FT W09719172-A1.
PD 29-MAY-1997..
PF 15-NOV-1996; J03356.
PR 30-NOV-1995; JP-311811.
PR 17-NOV-1995; JP-299611.
PA (ASAH) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR N-PSDB; T70175.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 5; Page 66-71; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1036 AA;
```

Query Match 33.1%; Score 574; DB 25; Length 1036;

Best Local Similarity 39.3%; Pred. No. 1.36e-48;

Matches 81; Conservative 47; Mismatches 70; Indels 8; Gaps 6;

```
Db 1 sgqfellellsmqnvngelqngccgarnpgdrktrdecctyfkvcikkeyqsvrtagpp 60
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 23 SGSGFELRLKYPFSDHGRDNEGRCCSGESDGTGKCL-GSCKTRFVCLKHQATIDTTSQ 81
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

- 61 csfsgstpvigntgnlfnkas-r-gn-d-prvlpfswfawprsytlilveav-dssnd-t 114
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
QY 82 CTYGDVITPILGNSVNLTDQRFQNGKFTNPQPFPSFSPGTFSLIVAWHDTNNSGN 141
  ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

Db 115 vqpdsi-iekashsgminpsrqwtlqntgvahfeyqirvtcdyyvgyfncnkfcrpr 173
  :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 142 ARTNKKLQRLVQVLEVSSEWTKNSQYTSLEYDFRVTCDLNYGSGCAFCRPRD 201
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 174 dffghyacdngnktcmegwmpcnc 199
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 202 DSFGHSTCSETGEIICLTGWGODYCH 227
  ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

RESULT 13

```
ID W18352 standard; protein; 1187 AA.
AC W18352;
DE 11-FEB-1998 (first entry)
KW Proliferation and differentiation suppression polypeptide.
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein /label= Signal
FT 32..1187 /label= Differentiation_suppression_protein
FT W09719172-A1.
PD 29-MAY-1997..
PF 15-NOV-1996; J03356.
PR 30-NOV-1995; JP-311811.
PR 17-NOV-1995; JP-299611.
PA (ASAH) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR N-PSDB; T70175.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 6; Page 71-76; 114pp; Japanese.
```

PR. 30-NOV-1995; JP-311811.

PR. 30-NOV-1995; JP-311811.

Molecular

(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 12 12:06:43 1998; Maspar time 2.40 Seconds
Tabular output not generated. 549.166 Million cell updates/sec

Title: >US-08-083-590A-2
Description: (1-230) from US08083590A.pep (2 of 2)
Perfect Score: 1734
Sequence: 1 MHWKILTAFTICTYIVQV.....ETGELICLTGMOGYCHTPK 230

Scoring table:
PAM 150
Gap 11

Searched: 62624 seqs, 5720858 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 30.190; Variance 125.451; scale 0.241

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 1734 | 100.0 | 833 | 1 | US-08-264-Sequence 6, Applicatio | 8.30e+164 |
| 2 | 1514 | 87.3 | 203 | 1 | US-08-264-Sequence 3, Applicatio | 3.57e+141 |
| 3 | 480 | 27.7 | 293 | 1 | US-08-264-Sequence 9, Applicatio | 3.16e+33 |
| 4 | 452 | 26.1 | 199 | 1 | US-08-264-Sequence 4, Applicatio | 1.84e+33 |
| 5 | 100 | 5.8 | 383 | 1 | US-08-457-Sequence 2, Applicatio | 1.11e+00 |
| 6 | 100 | 5.8 | 383 | 1 | US-08-597-Sequence 2, Applicatio | 1.11e+00 |
| 7 | 95 | 5.5 | 385 | 1 | US-08-597-Sequence 1, Applicatio | 2.73e+00 |
| 8 | 95 | 5.5 | 385 | 1 | US-08-457-Sequence 1, Applicatio | 2.73e+00 |
| 9 | 94 | 5.4 | 969 | 2 | PCT-US93-0-Sequence 2, Applicatio | 3.26e+00 |
| 10 | 92 | 5.3 | 768 | 1 | US-08-454-Sequence 4, Applicatio | 4.65e+00 |
| 11 | 92 | 5.3 | 769 | 1 | US-08-454-Sequence 6, Applicatio | 4.65e+00 |
| 12 | 88 | 5.1 | 316 | 1 | US-08-260-Sequence 18, Applicati | 9.36e+00 |
| 13 | 88 | 5.1 | 316 | 2 | PCT-US94-0-Sequence 18, Applicati | 9.36e+00 |
| 14 | 88 | 5.1 | 316 | 1 | US-08-017-Sequence 4, Applicatio | 1.11e+01 |
| 15 | 87 | 5.0 | 271 | 1 | US-08-447-Sequence 4, Applicatio | 1.11e+01 |
| 16 | 87 | 5.0 | 718 | 1 | US-08-444-Sequence 4, Applicatio | 1.11e+01 |
| 17 | 87 | 5.0 | 718 | 1 | US-08-445-Sequence 4, Applicatio | 1.11e+01 |
| 18 | 86 | 5.0 | 1810 | 2 | PCT-US95-1-Sequence 4, Applicatio | 1.32e+01 |
| 19 | 83 | 4.8 | 1124 | 2 | PCT-US93-0-Sequence 2, Applicatio | 2.21e+01 |
| 20 | 83 | 4.8 | 1124 | 1 | US-08-323-Sequence 2, Applicatio | 2.21e+01 |
| 21 | 83 | 4.8 | 1833 | 2 | PCT-US95-0-Sequence 18, Applicati | 2.21e+01 |
| 22 | 83 | 4.8 | 2199 | 2 | PCT-US95-1-Sequence 2, Applicatio | 2.21e+01 |
| 23 | 82 | 4.7 | 283 | 1 | US-08-658-Sequence 2, Applicatio | 2.61e+01 |

| RESULT | ID | US-08-264-534-6 | STANDARD: | PRT; | 833 AA. |
|--------|----|-----------------|-----------|------|-----------------------------------|
| 24 | 81 | 4.7 | 416 | 1 | US-08-252-Sequence 2, Applicatio |
| 25 | 81 | 4.7 | 464 | 1 | US-08-252-Sequence 6, Applicatio |
| 26 | 81 | 4.7 | 925 | 1 | US-08-252-Sequence 4, Applicatio |
| 27 | 81 | 4.7 | 1358 | 1 | US-08-404-Sequence 4, Applicatio |
| 28 | 81 | 4.7 | 1358 | 1 | US-08-404-Sequence 4, Applicatio |
| 29 | 81 | 4.7 | 1358 | 1 | US-08-404-Sequence 4, Applicatio |
| 30 | 79 | 4.6 | 650 | 1 | US-08-325-Sequence 56, Applicatio |
| 31 | 79 | 4.6 | 670 | 1 | US-08-477-Sequence 3, Applicatio |
| 32 | 79 | 4.6 | 670 | 1 | US-08-243-Sequence 3, Applicatio |
| 33 | 79 | 4.6 | 670 | 1 | US-08-484-Sequence 3, Applicatio |
| 34 | 79 | 4.6 | 688 | 1 | US-08-325-Sequence 57, Applicatio |
| 35 | 79 | 4.6 | 769 | 1 | US-08-484-Sequence 4, Applicatio |
| 36 | 79 | 4.6 | 769 | 1 | US-08-243-Sequence 4, Applicatio |
| 37 | 79 | 4.6 | 769 | 1 | US-08-477-Sequence 4, Applicatio |
| 38 | 78 | 4.5 | 501 | 1 | US-08-434-Sequence 2, Applicatio |
| 39 | 78 | 4.5 | 501 | 1 | US-07-687-Sequence 2, Applicatio |
| 40 | 78 | 4.5 | 501 | 1 | US-08-271-Sequence 2, Applicatio |
| 41 | 78 | 4.5 | 501 | 1 | PCT-US95-0-Sequence 2, Applicatio |
| 42 | 78 | 4.5 | 541 | 2 | PCT-US95-0-Sequence 30, Applicati |
| 43 | 77 | 4.4 | 47 | 1 | US-08-377-Sequence 2, Applicatio |
| 44 | 77 | 4.4 | 336 | 1 | US-08-093-Sequence 2, Applicatio |
| 45 | 77 | 4.4 | 336 | 1 | US-07-704-Sequence 3, Applicatio |

ALIGNMENTS

Sequence 6, Application US/08264534.
Sequence 6, Application US/08264534
Patent No. 5648464
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
TITLE OF INVENTION: In Toporychmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,534
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/695,189
FILING DATE: 03-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid


```

CC      MOLECULE TYPE:  protein
SQ      SEQUENCE 293 AA: 32286 MW: 438798 CN:

Query Match          27.7%;  Score 480;  DB 1;  Length 293;
Best Local Similarity 36.1%;  Pred.No.3,16-36;
Matches 82;  Conservative 59;  Mismatches 70;  Indels 16;  Gaps 11;

Db      63 CNLILMT-LILLVAKISAGNEFEILEISTNSHNLNGICCGPAELRAKTIIGSPCT 121
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      6 CLTFAFIEFTVYO-VHSGSFEELRLKFKNSDHRDNRDEGCCSGESDAGATGKCG-S-CR 62
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Db      122 TAFRLCEXYOTBEGASISTGCSFGNATTKIILGSSPVLSDP-----G-VGAIYLPFT 174
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      63 TFRFCLNHYQTLTD-IT-SQ-CYIGDVITPILSENSYNLTDAORFQNKFTNPDPFSS 119
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Db      175 FRWTKSFTLLIQ-A-LDMYNTISPPAERLIETSYSGVILP-SPEMKTLIDHIGRNARITVR 232
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      120 FSWPGETFLIYEAMHMDTNNSGNARTNKLILQLRLVQVLEVSSSEMKTKNSGESQYTSLEYD 179
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Db      233 VRVQCAVITYNTICTFPCRPDDQFGHAGCSSEQCKLCLNGWQSYNC 279
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
      180 FRVYTDLNYGSGCAKFCRPRDPSFGHSTCEIICLTGMOGDYC 226
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT      4          STANDARD:      PRT;      199 AA.
ID          US-08-264-534-4
XX          xxxxxx
AC          01-JAN-1900
DT          Sequence 4, Application US/08264534.
XX          Sequence 4, Application US/08264534.
DE          Sequence 4, Application US/08264534.
XX          Sequence 4, Application US/08264534
AC          Patent No. 5648464
CC          GENERAL INFORMATION:
CC          APPLICANT:  Aravanis-Tsakonas, Spyridon et al.
CC          TITLE OF INVENTION:  Human No. 5648464ch And Delta, Binding Domains
CC          TITLE OF INVENTION:  In Topotylmic Proteins, And Methods Based Thereon
CC          NUMBER OF SEQUENCES:  34
CC          CORRESPONDENCE ADDRESSES:
CC          ADDRESSEE:  Pennie & Edmonds
CC          STREET:  1155 Avenue of the Americas
CC          CITY:  New York
CC          STATE:  New York
CC          COUNTRY:  U.S.A.
CC          ZIP:  10036
CC          COMPUTER READABLE FORM:
CC          MEDIUM TYPE:  Floppy disk
CC          COMPUTER:  IBM PC compatible
CC          OPERATING SYSTEM:  PC-DOS/MS-DOS
CC          SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC          CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER:  US/08/264,534
CC          FILING DATE:
CC          CLASSIFICATION:  435
CC          PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER:  US 07/695,189
CC          FILING DATE:  03-MAY-1991
CC          ATTORNEY/AGENT INFORMATION:
CC          NAME:  Mistrock, S. Leslie
CC          REGISTRATION NUMBER:  18,872
CC          REFERENCE/DOCKET NUMBER:  7326-004
CC          TELECOMMUNICATION INFORMATION:
CC          TELEPHONE:  212 790-9090
CC          TELEFAX:  212 8698864/9741
CC          TELEX:  66141 PENNIE
CC          INFORMATION FOR SEQ ID NO: 4:
CC          SEQUENCE CHARACTERISTICS:
CC          LENGTH: 199 amino acids
CC          TYPE: amino acid
CC          STRANDEDNESS: single
CC          TOPOLOGY: unknown
CC

```

| | | |
|-------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| Dd | | 1 GNFEELEISTNSTNHLINGCCGPAELRAKTIGSPCTTARLCKEQTTEQCASTI 60 |
| Oy | | Best Local Similarity 37.0%; Pred. No. 1,84e-33; |
| Matches | 77; Conservative | 51; Mismatches 66; Indels 14; Gaps 9 |
| Dd | | 24 GFSEFLRKFGSDNRDRRCSCSGESDAGKCIQS-CRTRRVCLKHVOATID-TT- 79 |
| Oy | | 61 SNCGSFGNATTIKLGSSFSVLSDP-----G-VGAIVLPFRMTKSPTLLQA-LTMYN 112 : :::: :::: :::: :::: :::: :::: :::: :::: :::: |
| Dd | | 80 SQ-CYGDVITPILSENSVNLTDAORFONKGFTNPLOPFSSWPGTFSLVEAMHDTNN 138 |
| Oy | | 113 TSYPDALRIETISYSYGILP-SPEMKTLDIIGNRARITYVRVOCAYTYNTTCCTFCR 171 .: : :::: .: : :::: .: : :::: .: : :::: .: : :::: .: : :::: |
| Dd | | 139 SGNARTNKLIQLRLVVQVLEVSSWKTKNSGSQYTSLEYDFRYTCDLNYGGSCAFCR 198 |
| Oy | | 172 PRDDDPGRHACGSEGOKLCGLMGWGVCNC 199 : : : : 199 PRDDSFHGHSCTSETGEICLICLTGMOGDYC 226 |
| RESULT | 5 | STANDARD: |
| XX | US-08-457-135-2 | PRT: 383 AA. |
| AC | xxxxxx | |
| DT | 01-JAN-1900 | |
| Sequence 2, | Application US/08457135. | |
| XX | | |
| XX | | |
| Sequence 2, | Application US/08457135 | |
| CC | Patent No. 5644031 | |
| CC | GENERAL INFORMATION: | |
| CC | APPLICANT: LABORDA, Jorge | |
| CC | TITLE OF INVENTION: Delta-Like Gene Expressed In | |
| CC | TITLE OF INVENTION: Neuroendocrine Tumors | |
| CC | NUMBER OF SEQUENCES: 9 | |
| CC | CORRESPONDENCE ADDRESS: | |
| CC | ADDRESSEE: Foley & Lardner | |
| CC | STREET: 3000 K Street, N.W., Suite 500 | |
| CC | CITY: Washington, D.C. | |
| CC | COUNTRY: USA | |
| CC | ZIP: 20007-5109 | |
| CC | COMPUTER READABLE FORM: | |
| CC | MEDIUM TYPE: Floppy disk | |
| CC | COMPUTER: IBM PC compatible | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | |
| CC | SOFTWARE: Patentin Release #1.0, Version #1.25 | |
| CC | CURRENT APPLICATION DATA: | |
| CC | APPLICATION NUMBER: US/08/457,135 | |
| CC | FILING DATE: 01-JUN-1995 | |
| CC | PRIOR APPLICATION DATA: | |
| CC | APPLICATION NUMBER: 07/989,537 | |
| CC | FILING DATE: 11-DEC-1992 | |
| CC | ATTORNEY/AGENT INFORMATION: | |
| CC | NAME: BENT, Stephen A. | |
| CC | REGISTRATION NUMBER: 29,768 | |
| CC | REFERENCE/DOCKET NUMBER: 40359/304/NIID | |
| CC | TELECOMMUNICATION INFORMATION: | |
| CC | TELEPHONE: (202)672-5300 | |
| CC | TELEFAX: (202)672-5399 | |
| CC | TELEX: 904136 | |
| CC | INFORMATION FOR SEQ ID NO: 2: | |
| CC | SEQUENCE CHARACTERISTICS: | |
| CC | LENGTH: 383 amino acids | |
| CC | TYPE: amino acid | |
| CC | TOPOLOGY: linear | |
| CC | ORIGINAL SOURCE: | |
| CC | STRAIN: Human Dlx | |
| SEQ | SEQUENCE 383 AA: 41201 MW: 758273 CN: | |

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,135
CC FILING DATE: 01-JUN-1995
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/304/NIND
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 385 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
CC STRAIN: Mouse Dlx
CC SEQUENCE 385 AA; 41320 MW; 772804 CN;
SQ

Query Match 5.5%; Score 95; DB 1; Length 385;
Best Local Similarity 30.5%; Pred. No. 2,73e+00;
Matches 18; Conservative 9; Mismatches 30; Indels 2; Gaps 2;

Db 31 DPQYGFCEADNVCRCYHGWGGLCDK-CYTAFCV-NGVCKEPMQICIKDGMDFCEI 87
QY 170 ESQYTSLEVDFTCDLNYGSGCAKFCRPDRDSFGHSTCGEIGELICLTGMQGDYCHI 228

RESULT 9
ID PCT-US93-02147A-2 STANDARD; PRT; 969 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
XX
DE Sequence 2, Application PC/TUS9302147A.
XX
CC Sequence 2, Application PC/TUS9302147A.
CC GENERAL INFORMATION:
CC APPLICANT: BARR, PHILIP J
CC APPLICANT: KIEFER, MICHAEL C
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
CC TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
CC STREET: FIVE PALO ALTO SQUARE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/02147A

CC FILING DATE: 19930309
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/848,629
CC FILING DATE: 09-MAR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: NEELEY PH.D., RICHARD L
CC REGISTRATION NUMBER: 30092
CC REFERENCE/DOCKET NUMBER: CHIR-009/00US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 494-7622
CC TELEFAX: (415) 857-0663
CC TELEX: 380816 COOLEY PA
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 969 AA; 106419 MW; 4660240 CN;
SQ

Query Match 5.4%; Score 94; DB 2; Length 969;
Best Local Similarity 38.6%; Pred. No. 3,26e+00;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;

Db 733 CPLGYFDFTARRCR-RCHK-GCETCSSRAATQCLSCRGVYHH 774
QY 184 CDLNYGSGCAKFCRPDRDSFGHSTCGEIGELICLTGMQGDYCH 227

RESULT 10
ID US-08-454-455-4 STANDARD; PRT; 768 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
XX
DE Sequence 4, Application US/08454455.
XX
XX
CC Sequence 4, Application US/08454455
CC Patent No. 5635601
CC GENERAL INFORMATION:
CC APPLICANT: Moyle, Matthew
CC APPLICANT: McLean, John W.
CC TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.
CC STREET: 460 Point San Bruno Blvd
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WinPatIn (Genentech)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/454,455
CC FILING DATE: 30-MAY-1995
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/193989
CC FILING DATE: 09-FEB-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/004142
CC FILING DATE: 13-JAN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/670607
CC FILING DATE: 14-MAR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lee, Wendy M.

CC REGISTRATION NUMBER: 00,000
CC REFERENCE/DOCKET NUMBER: P0699C2D2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415/225-1994
CC TELEFAX: 415/952-9881
CC TELEX: 910/371-7168
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 768 amino acids
CC TYPE: Amino Acid
CC TOPOLOGY: Linear
SQ SEQUENCE 768 AA:84405 MW: 2945477 CN:

```
Query Match      5.3%; Score 92; DB 1; Length 768;
Best Local Similarity 48.0%; Pred. NO. 4.65e+00;
Matches 12; Conservative 6; Mismatches 6; Indels 1; Gaps 1;
```

| | | | |
|----|-----|------------------------------|-----|
| Db | 563 | GHGEC-EAGRCQCFSGWEGDRCQP | 586 |
| | | : : : : : : | |
| Qy | 205 | GHSTCSETGEIICLTGWQGDYCHIP | 229 |

| | | |
|--------|-----------------|--------------------------|
| RESULT | 11 | |
| ID | US-08-454-455-6 | STANDARD; - PRT; 769 AA. |
| XX | | |
| AC | XXXXXX | |
| XX | | |
| DT | 01 - JAN - 1900 | |

Sequence 6, Application US/08454455.
Sequence 6, Application US/08454455
Patent No. 5635601
GENERAL INFORMATION:
APPLICANT: Moyle, Matthew
APPLICANT: McLean, John W.
TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

| | | |
|----|--------------------------------------------------|--|
| CC | LENGTH: 769 amino acids | |
| CC | TYPE: Amino Acid | |
| CC | TOPOLOGY: Linear | |
| SQ | SEQUENCE 769 AA; 85631 MW; 2953052 CN; | |
| | Query Match 5.3%; Score 92; DB 1: Length 769; | |
| | Best Local Similarity 48.0%; Pred. No. 4.65e+00; | |
| | Matches 12; Conservative 6; Mismatches 6; Indels | |
| Db | 563 GHGEC-EAGRCQCFSGWGRDRCQP 586 | |
| QV | 205 GHSTCSETGETICITLWGQGDYCHIP 229 | |

```
Query Match          5.38; Score 92; DB 1; Length 769;
Best Local Similarity 48.0%; Pred. No. 4.65e+00;
Matches 12; Conservative 6; Mismatches 6; Indels 1; Gaps 1;
```

| | | | |
|----|-----|---------------------------|-----|
| Db | 563 | GHGEC-EAGRCQCFSGWEGDRCQP | 586 |
| | | : : :: : | |
| QY | 205 | GHSTCSETGEIICLTGWQGDYCHIP | 229 |

| | | | |
|--------|----------------------------------------|-----------|--------------|
| RESULT | 12 | | |
| ID | US-08-260-202A-18 | STANDARD; | PRT: 316 AA. |
| XX | | | |
| XX | xxxxxx | | |
| XX | | | |
| XX | | | |
| DT | 01-JAN-1900 | | |
| XX | | | |
| DE | Sequence 18; Application US/08260202A. | | |

Sequence 18, Application US/08260202A.

Sequence 18, Application US/08260202A

Patent No. 5573910

GENERAL INFORMATION:

APPLICANT: Deretic, Vojo

APPLICANT: Martin, Daniel W.

TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCCOIDY IN

TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENT

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P. O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/260,202A

Query Match 5.1%; Score 88; DB 1; Length 316;
Best Local Similarity 35.1%; Pred. No. 9.36e+00;
Matches 13: Conservative 9; Mismatches 13; Indels

db 38 NSFOGTFVYERNGSFS-THEIWHRVESDGAVR-ERLL 72

OY 112 NPIQPFSSWPGTFSLVEAMHDJNNSGNARTNKL 148

RESULT 13
ID PCT-US94-02034-18 STANDARD; PRT; 316 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 18, Application PC/TUS9402034.
XX
CC Sequence 18, Application PC/TUS9402034
CC GENERAL INFORMATION:
CC APPLICANT:
CC APPLICANT: STREET: 201 West 7th Street
CC APPLICANT: CITY: Austin
CC APPLICANT: STATE: Texas
CC APPLICANT: COUNTRY: United States of America
CC APPLICANT: POSTAL CODE: 78701
CC APPLICANT: TELEPHONE NO: (512)499-4462
CC APPLICANT: TELEFAX: (512)499-4523
CC TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN PSEUDOMONAS
CC TITLE OF INVENTION: AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P. O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: FLOPPY DISK
CC COMPUTER: IBM PC COMPATIBLE
CC OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/02034
CC FILING DATE: UNKNOWN
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/017,114
CC FILING DATE: 12 February 1993 (12.02.93)
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HODGINS, DANIEL S.
CC REGISTRATION NUMBER: 31,026
CC REFERENCE/DOCKET NUMBER: UTPK231P--
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 512/320-7200
CC TELEFAX: 713/789-2679
CC CC
CC TELEEX: 79-0924
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 316 AA; 34572 MW; 501312 CN;
DB 38 NSFOGTYEYERNGSFS-THEIWHRYESDGAIVR-ERLL 72
OY 112 NPIQPFSSWPGTFSLVEAMHDJNNSGNARTNKL 148
RESULT 14
ID US-08-017-114-18 STANDARD; PRT; 316 AA.
XX

AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 18, Application US/08017114.
XX
CC Sequence 18, Application US/08017114
CC Patent No. 5591838
CC GENERAL INFORMATION:
CC APPLICANT: Deretic, Vojto
CC APPLICANT: Martin, Daniel W.
CC TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCOIDY IN
CC TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENT
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P. O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/017,114
CC FILING DATE: 19930212
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hodgins, Daniel S.
CC REGISTRATION NUMBER: 31,026
CC REFERENCE/DOCKET NUMBER: UTSK:205/HOD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 512/320-7200
CC TELEFAX: 512/474-7577
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 316 AA; 34572 MW; 501312 CN;
DB 38 NSFOGTYEYERNGSFS-THEIWHRYESDGAIVR-ERLL 72
OY 112 NPIQPFSSWPGTFSLVEAMHDJNNSGNARTNKL 148
RESULT 15
ID US-08-447-554-4 STANDARD; PRT; 271 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 4, Application US/08447554.
XX
CC Sequence 4, Application US/08447554
CC Patent No. 5661003
CC GENERAL INFORMATION:
CC APPLICANT: FUSHIMI, KIYOHIIDE
CC APPLICANT: UCHIDA, SHINICHI
CC APPLICANT: SASAKI, SEI
CC APPLICANT: MARUMO, FUMIAKI
CC TITLE OF INVENTION: WATER CHANNEL

五、五、下、上、和、工

```
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
16-Feb-1997
ACCESSIONS A31246
REFERENCE A31246
#authors Kopcynski, C.C.; Alton, A.K.; Fichtel, K.; Kooh, P.J.;
#journal Muskavitch, M.A.T.
#title Genes Dev. (1988) 2:1723-1735
Delta, a Drosophila neurogenic gene, is transcriptionally
complex and encodes a protein related to blood coagulation
factors and epidermal growth factor of vertebrates.
#cross-references MUID:89196890
#accession A31246
#molecule_type mRNA
#residues 1-832 #label KOP
#cross-references GB:Y00222
GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463
#length 832 #molecular-weight 88943 #checksum 636
SUMMARY
Query Match 99.7%; Score 1728; DB 2; Length 832;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 229; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463
#length 832 #molecular-weight 88943 #checksum 636
SUMMARY
Query Match 99.7%; Score 1728; DB 2; Length 832;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 229; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 3
ENTRY S00670 #type complete
TITLE gene Delta protein precursor - fruit fly (Drosophila melanogaster)
ALTERNATE_NAMES gene D1 protein
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 16-Feb-1997
ACCESSIONS S00670
REFERENCE S00670
#authors Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
#journal EMBO J. (1987) 6:3431-3440
#title The neurogenic gene Delta of Drosophila melanogaster is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats.
#accession S00670
#molecule_type mRNA
#residues 1-880 #label VAE
#cross-references EMBL:X06289
GENETICS
#gene Delta
#cross-references FlyBase:FBgn0000463
#transmembrane protein
KEYWORDS #domain signal sequence #status predicted #label SIG\
1-18 #product gene Delta protein #status predicted #label MAT
19-88 #length 880 #molecular-weight 94643 #checksum 7597
SUMMARY
Query Match 99.3%; Score 1722; DB 2; Length 880;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 226; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGENSVNLTDAQRFQNGKFTNPQFPFSF 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 SWPGTSLIVEAWHDNNSGNARTNKLRLQVQVLEVSSEWKNKSESQYTSLEYDF 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGQGDYCHIPK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 5
ENTRY I48324 #type complete
TITLE DELTA-like 1 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997
ACCESSIONS I48324
REFERENCE I48324
#authors Bettenhausen, B.; de Angelis, M.H.; Simon, D.; Guenet, J.L.; Gossler, A.
#journal Development (1995) 121:2407-2418
#title Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta.
```



```
##cross-references FlyBase:FBgn0004197
CLASSIFICATION #superfamily EGF homology
KEYWORDS glycoprotein; transmembrane protein
FEATURE
1-84
85-1408
#domain signal sequence #status predicted #label sig\
#product gene serrate protein #status predicted #label
MAT\
#domain extracellular #status predicted #label EXT\
#domain EGF homology #label EG01\
#domain EGF homology #label EG02\
#domain EGF homology #label EG03\
#domain EGF homology #status atypical #label EG04\
#domain EGF homology #label EG05\
#domain EGF homology #status atypical #label EG06\
#domain EGF homology #label EG07\
#domain EGF homology #label EG08\
#domain EGF homology #label EG09\
#domain EGF homology #status atypical #label EG10\
#domain EGF homology #label EG11\
#domain EGF homology #label EG12\
#domain EGF homology #label EG13\
#domain EGF homology #label EG14\
#region cysteine-rich
#domain transmembrane #status predicted #label TM1\
#domain intracellular #status predicted #label INT\
1222-1246
1247-1408
152,196,247,331,
412,452,558,739,
965,977,1004,1030,
1150
#binding_site carbohydrate (Asn) (covalent) #status
predicted
#length 1408 #molecular-weight 150645 #checksum 5466
SUMMARY
Query Match 27.7%; Score 480; DB 2; Length 1408;
Best Local Similarity 36.1%; Pred. No. 2.63e-76;
Matches 82; Conservative 59; Mismatches 70; Indels 16; Gaps 11;

Db 67 CNLIALI-LILLVHKISAGNFLEILEISNTNSHLLNGYCCGPAELRATKTIKCSPT 125
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 6 CLLTAFICFTVIVQ-VHSSGSFELRLKFNDRHNGRCCSGESDGTGKCLG-S-CK 62
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 126 TAFRLCLKEYQTTEQASISITGCSFGNATTKILGGSFVLSDP-----G-VGAIVLPFT 178
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 TRFVCLKHVQATID-TT-SQ-CTYGDVITPILGENSVNLTAQRFQNKFTNPQFPFS 119
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 FRWTKSFTILQALDMYNTSYDPAERLIEETSYSGVILP-SPEWKYLDHIGRNARTYR 236
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 FSWPGTFSLIVEAWHDNTNSGNARTKLLIQLLVQOVLEVSEWTKNKSSEQVTSLEYD 179
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 VRVQCAVYNTTCTTFCRPDPOFGHYACSGEGOKLCLNGWGVNC 283
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 FRVTCNLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYC 226
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
ENTRY #type complete
TITLE lag-2 protein - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
09-Sep-1997
ACCESSIONS S42367
REFERENCE #status preliminary
#authors Tax, F.E.; Yeagers, J.J.; Thomas, J.H.
#journal Nature (1994) 368:150-154
#title Sequence of C. elegans lag-2 reveals a cell-signalling domain
shared with Delta and Serrate of Drosophila.
#accession S42367
#status preliminary
#molecule_type DNA
#residues 1-402 #label TAX
#cross-references EMBL:X77495; NID:g459578; PID:g459579
GENETICS
#introns 168/3; 241/3
SUMMARY #length 402 #molecular-weight 44603 #checksum 6894
```

```
Query Match 7.7%; Score 133; DB 2; Length 402;
Best Local Similarity 35.6%; Pred. No. 7.32e-07;
Matches 16; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Db 122 VTCARNYFGNRCNFCDAHLAKAKRCDAMGRLCRDIGWMPHC 166
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 VTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYC 226
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 10
ENTRY #type complete
TITLE delta-like dlk homeotic protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Sep-1993 #sequence_revision 22-Apr-1995 #text_change
10-Sep-1997
ACCESSIONS B45484; S31973; S31974
REFERENCE A45484
#authors Laborde, J.; Sausville, E.A.; Hoffman, T.; Notario, V.
#journal J. Biol. Chem. (1993) 268:3817-3820
#title dlk, a putative mammalian homeotic gene differentially
expressed in small cell lung carcinoma and neuroendocrine
tumor cell line.
#cross-references MUID:93179372
#accession B45484
#molecule_type mRNA
#residues 1-383 #label LAB
#cross-references EMBL:Z12172; NID:g38478; PID:g38479
#note sequence extracted from NCBI backbone (NCBI:125735)
#note species designations for this sequence report and for
A45484 originally were transposed in NCBI backbone
SUMMARY #length 383 #molecular-weight 41154 #checksum 4637
```

```
Query Match 6.0%; Score 104; DB 2; Length 383;
Best Local Similarity 32.6%; Pred. No. 2.36e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;

Db 45 CHVGWQPLCDQ-CVTSFGCL-HGLGEGQCICITDGDGELC 85
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 CDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYC 226
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 11
ENTRY #type complete
TITLE integrin beta-3 subunit - African clawed frog
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
16-Feb-1997
ACCESSIONS I51530
REFERENCE I51530
#authors Ransom, D.G.; Hens, M.D.; DeSimone, D.W.
#journal Dev. Biol. (1993) 160:265-275
#title Integrin expression in early amphibian embryos: cDNA cloning
and characterization of Xenopus beta 1, beta 2, beta 3, and
beta 6 subunits.
#cross-references MUID:94040374
#accession I51530
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-788 #label RAN
#cross-references GB:L13591; NID:g467811; PID:g467812
CLASSIFICATION #superfamily integrin beta chain
SUMMARY #length 788 #molecular-weight 87490 #checksum 7411
```

```
Query Match 5.9%; Score 103; DB 2; Length 788;
Best Local Similarity 39.5%; Pred. No. 3.29e-02;
Matches 15; Conservative 7; Mismatches 12; Indels 4; Gaps 4;

Db 547 CDDFSLRYKGCMSGHQCN-CGDCICESDWTGYCN 583
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 CAKF-C-RPRDD-SFGHSTCSETGEIICLTGWQGDYCH 227
| | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
```


RESULT 12
ENTRY S48713 #type complete
TITLE fetal antigen 1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 01-Aug-1995 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
ACCESSIONS S48713
REFERENCE S48713
#authors Jensen, C.H.; Krogh, T.N.; Hojrup, P.; Clausen, P.P.; Skjodt, K.; Larsson, L.I.; Englund, J.J.; Teisner, B.
#journal Eur. J. Biochem. (1994) 225:83-92
#title Protein structure of fetal antigen 1 (FA1). A novel circulating human epidermal-growth-factor-like protein expressed in neuroendocrine tumors and its relation to the gene products of dlk and pg2.
#accession S48713
#status preliminary
#molecule_type protein
#residues 1-259 #label JEN
SUMMARY #length 259 #molecular_weight 27316 #checksum 4381
Query Match 5.8%; Score 100; DB 2; Length 259;
Best Local Similarity 32.6%; Pred. No. 8.83e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;
Db 22 CQPGWGPICDQ-CVTSFGCL-HGICGEPGOCICTDGMDELC 62
QY 184 CDLNYGSGCAKFCRPRDSDFGHSTCSETGEIICLTGMOGDYC 226
RESULT 13
ENTRY A44549 #type complete
TITLE fetal antigen 1 homeotic protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Feb-1994
ACCESSIONS A44549
REFERENCE A44549
#authors Hojrup, P.; Jensen, C.H.; Skjodt, K.; Teisner, B.
#journal Protein Sci. (1993) 2(Suppl.1):2595
#title Primary structure of human fetal antigen 1 (FA1), a putative homeotic glycoprotein.
#accession A44549
#status preliminary
#molecule_type protein
#residues 1-260 #label HOJ
SUMMARY #length 260 #molecular_weight 27413 #checksum 7173
Query Match 5.8%; Score 100; DB 2; Length 260;
Best Local Similarity 32.6%; Pred. No. 8.83e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;
Db 22 CQPGWGPICDQ-CVTSFGCL-HGICGEPGOCICTDGMDELC 62
QY 184 CDLNYGSGCAKFCRPRDSDFGHSTCSETGEIICLTGMOGDYC 226
RESULT 14
ENTRY S53716 #type complete
TITLE homeotic protein dlk long form - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 10-Sep-1997
ACCESSIONS S53716; S53717; S71068
REFERENCE S53716
#authors Lee, Y.L.; Helman, L.; Hoffman, T.; Laborda, J.
#journal Biochim. Biophys. Acta (1995) 1261:223-232
#title dlk, pg2 and Pref-1 mRNAs encode similar proteins belonging to the EGF-like superfamily. Identification of polymorphic variants of this RNA.
#accession S53716
#molecule_type mRNA

##residues 1-383 #label LEE
##cross-references EMBL:U15979
#note the authors translated the codon CAG for residue 46 as His and CCT for residue 47 as Val
#accession S53717
#molecule_type mRNA
#residues 1-228,302-383 #label LEM
##cross-references EMBL:U15981; NID:9562109; PID:9562110
REFERENCE S71068
#authors Laborda, J.
#submission submitted to the EMBL Data Library, October 1994
#accession S71068
#molecule_type mRNA
#residues 1-344,346-383 #label LAB
##cross-references EMBL:U15979; NID:9562105; PID:9562106
GENETICS
#gene dlk
#keywords alternative splicing; transmembrane protein
FEATURE
1-383
1-228,302-383 #product homeotic protein dlk long form #status predicted #label MAT1
SUMMARY #length 383 #molecular_weight 41143 #checksum 4769
Query Match 5.8%; Score 100; DB 2; Length 383;
Best Local Similarity 32.6%; Pred. No. 8.83e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;
Db 45 CQPGWGPICDQ-CVTSFGCL-HGICGEPGOCICTDGMDELC 85
QY 184 CDLNYGSGCAKFCRPRDSDFGHSTCSETGEIICLTGMOGDYC 226
RESULT 15
ENTRY A37057 #type complete
TITLE integrin beta-6 chain - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 31-Oct-1997
ACCESSIONS A37057
REFERENCE A37057
#authors Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pyeleta, R.
#journal J. Biol. Chem. (1990) 265:11502-11507
#title Complete amino acid sequence of a novel integrin beta subunit (beta6) identified in epithelial cells using the polymerase chain reaction.
#cross-references MUID:90307659
#accession A37057
#status preliminary
#molecule_type mRNA
#residues 1-788 #label SHE
##cross-references GB:M35198; GB:J05522; NID:9186506; PID:9186507
CLASSIFICATION #superfamily Integrin beta chain
KEYWORDS cell adhesion; cytoskeleton; transmembrane protein
SUMMARY #length 788 #molecular_weight 85975 #checksum 8091
Query Match 5.8%; Score 101; DB 2; Length 788;
Best Local Similarity 36.8%; Pred. No. 6.37e-02;
Matches 14; Conservative 9; Mismatches 11; Indels 4; Gaps 4;
Db 539 CDFSCVHRHKGILGNGDC-DCGECVCRSGWTGECN 575
QY 193 CAKF-C-RPRD-SFGHSTCSETGEIICLTGMOGDYCH 227
Search completed: Fri Jun 12 12:04:35 1998
Job time : 39 secs.

NWSELE
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 12 12:04:53 1998; MasPar time 7.25 Seconds
Tabular output not generated. 795.610 Million cell updates/sec

Title: >US-08-083-590A-2
Description: (1-230) from US08083590A.pep (2 of 2)
Perfect Score: 1734
Sequence: 1 MHMVKCLTFAICTFTIVQV.....ETGETICLTGMGDCYCHPK 230

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 44.483; Variance 66.618; scale 0.668

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 1722 | 99.3 | 880 | 1 | DL_DROME NEUROGENIC LOCUS DELTA | 0.00e+00 |
| 2 | 732 | 42.2 | 722 | 1 | DL_MOUSE DELTA-LIKE PROTEIN 1 P | 1.85e-161 |
| 3 | 726 | 41.9 | 714 | 1 | DL_MOUSE DELTA-LIKE PROTEIN 1 P | 8.02e-160 |
| 4 | 480 | 27.7 | 1408 | 1 | SERRA_RAT SERRATE PROTEIN PRECUR | 1.25e-93 |
| 5 | 133 | 7.7 | 402 | 1 | LAG2_MOUSE LAG-2 PROTEIN PRECURSO | 7.12e-09 |
| 6 | 116 | 6.7 | 515 | 1 | APX1_MOUSE APX-1 PROTEIN PRECURSO | 1.28e-05 |
| 7 | 100 | 5.8 | 383 | 1 | DL_MOUSE DELTA-LIKE PROTEIN PRE | 8.66e-03 |
| 8 | 101 | 5.8 | 788 | 1 | ITB6_HUMAN INTEGRIN BETA-6 SUBUNI | 5.87e-03 |
| 9 | 97 | 5.6 | 301 | 1 | CD42_YEAST CHITIN DEACETYLASE 2 P | 2.74e-02 |
| 10 | 97 | 5.6 | 577 | 1 | ITB6_MOUSE INTEGRIN BETA-6 SUBUNI | 2.74e-02 |
| 11 | 97 | 5.6 | 3672 | 1 | LM2_MOUSE LAMININ-LIKE PROTEIN K | 2.74e-02 |
| 12 | 95 | 5.5 | 385 | 1 | DL_MOUSE DELTA-LIKE PROTEIN PRE | 5.81e-02 |
| 13 | 94 | 5.4 | 810 | 1 | YL25_MOUSE HYPOTHETICAL 89.2 KD P | 8.43e-02 |
| 14 | 94 | 5.4 | 969 | 1 | PAC4_HUMAN SUBTILISIN-LIKE PROTEA | 8.43e-02 |
| 15 | 92 | 5.3 | 590 | 1 | CAN1_YEAST ARGININE PERMEASE | 1.76e-01 |
| 16 | 92 | 5.3 | 768 | 1 | ITB8_MOUSE INTEGRIN BETA-8 SUBUNI | 1.76e-01 |
| 17 | 92 | 5.3 | 769 | 1 | ITB8_HUMAN INTEGRIN BETA-8 SUBUNI | 1.76e-01 |
| 18 | 90 | 5.2 | 937 | 1 | CP5B_MOUSE CYTOCHROME P450 LIT1A2 | 3.62e-01 |
| 19 | 90 | 5.2 | 937 | 1 | PAC4_RAT SUBTILISIN-LIKE PROTEA | 3.62e-01 |
| 20 | 91 | 5.2 | 3610 | 1 | LMAS_MOUSE LAMININ ALPHA-5 CHAIN | 2.53e-01 |
| 21 | 90 | 5.2 | 4540 | 1 | DYHC_MOUSE DYENIN HEAVY CHAIN | 3.62e-01 |
| 22 | 88 | 5.1 | 47 | 1 | THG2_WHEAT GAMMA-2 PURIFICATIONIN | 7.35e-01 |
| 23 | 89 | 5.1 | 238 | 1 | UBC3_RABIT UBIQUITIN-CONJUGATING | 5.17e-01 |

| | | | | | | |
|----|----|-----|------|---|-----------------------------------|----------|
| 24 | 88 | 5.1 | 316 | 1 | MUCB_PSEAE SIGMA FACTOR ALGU REGU | 7.35e-01 |
| 25 | 87 | 5.0 | 271 | 1 | AOP2_RAT AQUOPORIN-CD (AOP-CD) | 1.04e+00 |
| 26 | 87 | 5.0 | 298 | 1 | YAKI_YEAST HYPOTHETICAL 35.1 KD P | 1.04e+00 |
| 27 | 87 | 5.0 | 788 | 1 | ITB3_HUMAN PLATELET MEMBRANE GLYC | 1.04e+00 |
| 28 | 86 | 5.0 | 858 | 1 | CVAG_DICDI ADENYLATE CYCLASE, GER | 1.47e+00 |
| 29 | 87 | 5.0 | 1295 | 1 | GLP1_MOUSE GLP-1 PROTEIN PRECURSO | 1.04e+00 |
| 30 | 86 | 5.0 | 1808 | 1 | TENA_CHICK TENASCIN PRECURSOR (TN | 1.47e+00 |
| 31 | 85 | 4.9 | 543 | 1 | PAT2_YEAST PEROXISOMAL-COENZYME A | 2.08e+00 |
| 32 | 85 | 4.9 | 672 | 1 | YCDR_ECOLI HYPOTHETICAL 77.4 KD L | 2.08e+00 |
| 33 | 85 | 4.9 | 3649 | 1 | ACVS_MOUSE DELTA-(L)-ALPHA-AMINOD | 2.08e+00 |
| 34 | 84 | 4.8 | 87 | 1 | Y174_BURCE INSERTION ELEMENT IS40 | 2.91e+00 |
| 35 | 83 | 4.8 | 284 | 1 | YKGD_ECOLI HYPOTHETICAL TRANSCRIP | 4.08e+00 |
| 36 | 83 | 4.8 | 328 | 1 | CHIT_PHAVU ENDOCHITINASE PRECURSO | 4.08e+00 |
| 37 | 84 | 4.8 | 396 | 1 | DUT_VZWD DEOXYRIBIDINE 5'-TRIPHOS | 2.91e+00 |
| 38 | 84 | 4.8 | 433 | 1 | YB14_YEAST HYPOTHETICAL 50.8 KD P | 2.91e+00 |
| 39 | 83 | 4.8 | 456 | 1 | YB8_SCHPO PUTATIVE 73.4 KD TRANS | 4.08e+00 |
| 40 | 83 | 4.8 | 637 | 1 | YAO7_SCHPO FASCICLIN II, PHOSPHAT | 4.08e+00 |
| 41 | 83 | 4.8 | 811 | 1 | FS22_DROME FASCICLIN II, MEMBRANE | 4.08e+00 |
| 42 | 83 | 4.8 | 873 | 1 | TIE2_DROME ANGIOPOIETIN 1 RECEPTO | 4.08e+00 |
| 43 | 83 | 4.8 | 1124 | 1 | TIE2_HUMAN TENASCIN PRECURSOR (TN | 4.08e+00 |
| 44 | 83 | 4.8 | 2199 | 1 | PCCA_HUMAN AGGRECAN CORE PROTEIN | 4.08e+00 |
| 45 | 83 | 4.8 | 2415 | 1 | | |

ALIGNMENTS

RESULT 1
ID DL_DROME STANDARD; PRT; 880 AA.
AC P10041;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.
GN DL.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A.
RA VAESSIN H., BREMER K.A., KNUST E., CAMPOS-ORTEGA J.A.;
RL EMBO J. 6:3431-3440(1987).
RN [2]
RP SEQUENCE OF 422-621 FROM N.A.
RX MEDLINE; 87218537.
RA KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D., VAESSIN H.,
RL CAMPOS-ORTEGA J.A.;
RN EMBO J. 6:761-766(1987).
RN [3]
RP PATTERN OF TRANSCRIPTION.
RX MEDLINE; 91209246.
RA HENSLIN M., KRANATSCHEK B., CAMPOS-ORTEGA J.A.;
RL DEVELOPMENT 110:905-914(1990).
CC -1- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL
IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL
CELL LINESAGES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART
OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS.
CC THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
CC -1- NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL. IT IS
CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPLETE
FOR BINDING WITH THE NOTCH PROTEIN.
CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: TO DROSOPHILA SERRATE PROTEIN.
DR EMBL; X06289; G7853; -.
DR EMBL; X05140; G929563; -.
DR PIR; S00670; S00670.
DR PIR; A26637; A26637.
DR HSSP; P00725; ICBH.
DR FLYBASE; FBgn000463; DL.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 2.

KW DIFFERENTIATION; NEUROGENESIS; REPEAT; TRANSMEMBRANE;
EGF-LIKE DOMAIN; GLYCOPROTEIN; SIGNAL;
FT SIGNAL 1 18
FT CHAIN 19 880
FT DOMAIN 19 653
FT TRANSMEM 654 677
FT DOMAIN 678 880
FT DOMAIN 227 258
FT DOMAIN 256 289
FT DOMAIN 291 329
FT DOMAIN 331 372
FT DOMAIN 374 416
FT DOMAIN 418 451
FT DOMAIN 453 489
FT DOMAIN 491 527
FT DOMAIN 529 565
FT DISULFID 231 240
FT DISULFID 235 246
FT DISULFID 248 257
FT DISULFID 260 271
FT DISULFID 266 277
FT DISULFID 279 288
FT DISULFID 295 307
FT DISULFID 301 317
FT DISULFID 319 328
FT DISULFID 335 348
FT DISULFID 342 360
FT DISULFID 362 371
FT DISULFID 378 388
FT DISULFID 383 404
FT DISULFID 406 415
FT DISULFID 422 433
FT DISULFID 427 439
FT DISULFID 441 450
FT DISULFID 457 468
FT DISULFID 462 477
FT DISULFID 479 488
FT DISULFID 495 506
FT DISULFID 500 515
FT DISULFID 517 526
FT DISULFID 533 544
FT DISULFID 538 553
FT DISULFID 555 564
FT CARBOHYD 98 98
FT CARBOHYD 137 137
FT CARBOHYD 167 167
FT CARBOHYD 649 649
FT CONFLICT 437 438
FT CONFLICT 459 459
FT CONFLICT 490 490
SQ SEQUENCE 880 AA; 94643 MW; E967E862 CRC32;

Query Match 99.3%; Score 1722; DB 1; Length 880;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 228; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWINCLTAFCTVIVOVHSSGSGFELRLKYFSDNHDGRNRCGSGSDGATGKCLGS 60
Qy 1 MHWIKLLTAFCTVIVOVHSSGSGFELRLKYFSDNHDGRNRCGSGSDGATGKCLGS 60
Db 61 CKTERLCCLKHYQATIDTTSQCTYGDVITPILGENSVNLTDQRFONKGFNPQPPFSF 120
Qy 61 CKTERVCLKHQATIDTTSQCTYGDVITPILGENSVNLTDQRFONKGFNPQPPFSF 120
Db 121 SWPGFSLIVEAHHWTNNSGNARTNKLQRLVQVLEVSSEWTKNKSQVTSLEYDF 180
Qy 121 SWPGFSLIVEAHHWTNNSGNARTNKLQRLVQVLEVSSEWTKNKSQVTSLEYDF 180
Db 181 RVTCDLNTYGGCAKFCRPRDSDFGHSTCSETGEIICLTGWOQDYCHIPK 230
Qy 181 RVTCDLNTYGGCAKFCRPRDSDFGHSTCSETGEIICLTGWOQDYCHIPK 230

RESULT 2
ID DLL1 MOUSE STANDARD; PRT; 722 AA.
AC 061483;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA1).
GN DLL1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C X C57BL/6; TISSUE-EMBRYO;
RX MEDLINE; 95401858.
RA BETTENHAUSEN B., DE ANGELIS M.H., SIMON D., GUENET J.-L., GOSSLER A.;
RL DEVELOPMENT 121:2407-2418(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN ADULT
MESODERM AND NERVOUS SYSTEM. EXPRESSED AT HIGH LEVELS IN ADULT
HEART AND AT LOWER LEVELS, IN ADULT LUNG.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED UNTIL DAY 15 IN THE EMBRYO.
CC EXPRESSION THEN DECREASES AND INCREASES AGAIN IN THE ADULT.
CC -!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: TO DROSOPHILA DELTA PROTEIN.
DR EMBL; X80903; G806570;
DR MGD; MGI:104659; DLL1.
DR PROSITE; PS00010; ASX HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 2.
KW SIGNAL; EGF-LIKE DOMAIN; GLYCOPROTEIN; TRANSMEMBRANE.
FT SIGNAL 1 17
FT CHAIN 18 722
FT DOMAIN 18 545
FT TRANSMEM 546 588
FT DOMAIN 569 722
FT DOMAIN 225 253
FT DOMAIN 256 284
FT DOMAIN 291 324
FT DOMAIN 321 362
FT DOMAIN 369 401
FT DOMAIN 408 439
FT DOMAIN 445 477
FT DOMAIN 484 515
FT DOMAIN 515 526
FT DISULFID 225 236
FT DISULFID 229 242
FT DISULFID 244 253
FT DISULFID 256 267
FT DISULFID 262 273
FT DISULFID 275 284
FT DISULFID 291 303
FT DISULFID 297 313
FT DISULFID 315 324
FT DISULFID 331 342
FT DISULFID 336 351
FT DISULFID 353 362
FT DISULFID 369 380
FT DISULFID 374 390
FT DISULFID 392 401
FT DISULFID 408 419
FT DISULFID 413 428
FT DISULFID 430 439
FT DISULFID 446 456
FT DISULFID 468 477
FT DISULFID 484 495
FT DISULFID 489 504
FT DISULFID 506 515
FT CARBOHYD 476 476
SQ SEQUENCE 722 AA; 78448 MW; 5A647702 CRC32;

| ID | APX1 CAEEL | STANDARD: | PRT: | 515 AA. |
|------|--------------------------------------------------------------------|-----------|------|---------|
| AC | P41990: | | | |
| AD | 01-NOV-1995 (REL. 32, CREATED) | | | |
| AE | 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE) | | | |
| AF | 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE) | | | |
| AG | APX-1: PROTEIN PRECURSOR. | | | |
| AH | APX-1: | | | |
| AI | CARONABDITIS ELEGANS. | | | |
| AJ | EDUAROTA: METACOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA. | | | |
| AK | [1] | | | |
| AL | SEQUENCE FROM N.A. | | | |
| AM | STRAIN-BRISTOL N2. | | | |
| AN | MEDLINE: 94208066. | | | |
| AO | MELLO C.C., DRAPER B.W., PRIEST J.R.: | | | |
| AP | CELL 77:95-106(1994). | | | |
| AQ | -1- FUNCTION: INVOLVED IN THE SPECIFICATION OF THE BLASTOMER CELL | | | |
| AR | AP FATE. CONTRIBUTES TO THE ESTABLISHMENT OF THE DORSAL-VENTRAL | | | |
| AS | AXIS IN THE EARLY C.ELEGANS EMBRYO. | | | |
| AT | -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN. | | | |
| AV | -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN. | | | |
| AW | EMBL: G47381.1; - | | | |
| AX | PROSITE: PS00022; EGF_1; 4. | | | |
| AY | PROSITE: PS01186; EGF_2; 3. | | | |
| AZ | DIFFERENTIATION: REPEAT; TRANSMEMBRANE; EGF-LIKE DOMAIN; | | | |
| BA | GLYCOPROTEIN: SIGNAL. | | | |
| BB | SIGNAL | | | |
| BC | CHAIN | | | |
| BD | DOMAIN | | | |
| BE | DOMAIN | | | |
| BF | TRANSMEM | | | |
| BG | DOMAIN | | | |
| BH | DOMAIN | | | |
| BI | DOMAIN | | | |
| BJ | DOMAIN | | | |
| BK | DOMAIN | | | |
| BL | DOMAIN | | | |
| BM | DOMAIN | | | |
| BN | DOMAIN | | | |
| BO | DOMAIN | | | |
| BP | DOMAIN | | | |
| BQ | DOMAIN | | | |
| BR | DOMAIN | | | |
| BS | DOMAIN | | | |
| BT | DOMAIN | | | |
| BU | DOMAIN | | | |
| BV | DOMAIN | | | |
| BW | DOMAIN | | | |
| BY | DOMAIN | | | |
| BZ | DOMAIN | | | |
| CA | DOMAIN | | | |
| CB | DOMAIN | | | |
| CC | DOMAIN | | | |
| CD | DOMAIN | | | |
| CE | DOMAIN | | | |
| CF | DOMAIN | | | |
| CG | DOMAIN | | | |
| CH | DOMAIN | | | |
| CI | DOMAIN | | | |
| CJ | DOMAIN | | | |
| CK | DOMAIN | | | |
| CL | DOMAIN | | | |
| CM | DOMAIN | | | |
| CN | DOMAIN | | | |
| CO | DOMAIN | | | |
| CP | DOMAIN | | | |
| CQ | DOMAIN | | | |
| CR | DOMAIN | | | |
| CS | DOMAIN | | | |
| CT | DOMAIN | | | |
| CU | DOMAIN | | | |
| CV | DOMAIN | | | |
| CW | DOMAIN | | | |
| CX | DOMAIN | | | |
| CY | DOMAIN | | | |
| CA | DOMAIN | | | |
| CB | DOMAIN | | | |
| CC | DOMAIN | | | |
| CD | DOMAIN | | | |
| CE | DOMAIN | | | |
| CF | DOMAIN | | | |
| CG | DOMAIN | | | |
| CH | DOMAIN | | | |
| CI | DOMAIN | | | |
| CJ | DOMAIN | | | |
| CK | DOMAIN | | | |
| CL | DOMAIN | | | |
| CM | DOMAIN | | | |
| CN | DOMAIN | | | |
| CO | DOMAIN | | | |
| CP | DOMAIN | | | |
| CQ | DOMAIN | | | |
| CR | DOMAIN | | | |
| CS | DOMAIN | | | |
| CT | DOMAIN | | | |
| CU | DOMAIN | | | |
| CV | DOMAIN | | | |
| CW | DOMAIN | | | |
| CX | DOMAIN | | | |
| CY | DOMAIN | | | |
| CA | DOMAIN | | | |
| CB | DOMAIN | | | |
| CC | DOMAIN | | | |
| CD | DOMAIN | | | |
| CE | DOMAIN | | | |
| CF | DOMAIN | | | |
| CG | DOMAIN | | | |
| CH | DOMAIN | | | |
| CI | DOMAIN | | | |
| CJ | DOMAIN | | | |
| CK | DOMAIN | | | |
| CL | DOMAIN | | | |
| CM | DOMAIN | | | |
| CN | DOMAIN | | | |
| CO | DOMAIN | | | |
| CP | DOMAIN | | | |
| CQ | DOMAIN | | | |
| CR | DOMAIN | | | |
| CS | DOMAIN | | | |
| CT | DOMAIN | | | |
| CU | DOMAIN | | | |
| CV | DOMAIN | | | |
| CW | DOMAIN | | | |
| CX | DOMAIN | | | |
| CY | DOMAIN | | | |
| CA | DOMAIN | | | |
| CB</ | | | | |

[illegible]

| | |
|----|-------------------------------------------------------------------------|
| RA | MISHRA C.V., SEMINO G.E., MCCREATH K.J., DE LA VEGA H., JONES B.J., |
| RA | SPECHT C.A., ROBBINS P.W.: |
| RL | YEAST 13:327-336(1987). |
| CC | -I- FUNCTION: HYDROLYSES THE N-ACETAMIDO GROUPS OF N-ACETYL-D- |
| CC | GLUCOSAMINE RESIDUES IN CHITIN. |
| CC | -I- CATALYTIC ACTIVITY: CHITIN + H ₂ O = CHITOSAN + ACETATE. |
| CC | -I- DEVELOPMENTAL STAGE: EXPRESSED EXCLUSIVELY DURING SPOULATION. |
| CC | -I- SIMILARITY: TO OTHER POLYSACCHARIDE DEACETYLASES. |
| DR | EMBL; U17247; G577218; - |
| DR | SCD; L0003113; CDA2. |
| KW | HYDROLASE; CHITIN DEGRADATION; SIGNAL; GLYCOPROTEIN; SPOULATION. |
| FT | SIGNAL |
| FT | CHAIN 1 ? POTENTIAL. |
| FT | CARBOHYD 301 CHITIN DEACETYLASE 2. |
| FT | CARBOHYD 26 POTENTIAL. |
| FT | CARBOHYD 50 POTENTIAL. |
| FT | CARBOHYD 68 POTENTIAL. |
| FT | CARBOHYD 189 POTENTIAL. |
| FO | SEQUENCE 301 AA; 34642 MW; 1759E4C7 CRC32; |

[illegible]

| RESULT | ID | 10 | STANDARD; | PRT; | 577 | AA. |
|--------|---------------------------------------------------------------------|-----------------------------------|-----------|------|-----|-----|
| AC | ITB6 | CAVPO | | | | |
| AD | P18563; | | | | | |
| DT | 01-NOV-1990 | (REL. 16, CREATED) | | | | |
| DT | 01-NOV-1990 | (REL. 16, LAST SEQUENCE UPDATE) | | | | |
| DT | 01-OCT-1996 | (REL. 34, LAST ANNOTATION UPDATE) | | | | |
| DE | INTEGRIN BETA-6 | SUBUNIT (FRAGMENT). | | | | |
| GN | ITGB6. | | | | | |
| OS | CAVIA PORCELLUS (GUINEA PIG). | | | | | |
| OC | EUMAROTIA, METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; | | | | | |
| OC | EUTHERIA; RODENTIA. | | | | | |
| RN | [1] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | |
| RC | STRAIN-HARTLEY; | | | | | |
| RX | MEDLINE; 90307659. | | | | | |
| RA | SHEPPARD D., ROZZO C., STARR L., QUARANTA V., ERLE D.J., PTELA R.; | | | | | |
| RL | J. BIOL. CHEM. 265:11502-11507(1990) | | | | | |
| CC | -1- FUNCTION: INTEGRIN ARE A LARGE FAMILY OF CELL SURFACE | | | | | |
| CC | GLYCOPROTEINS THAT MEDIATE CELL TO CELL & CELL TO MATRIX ADHESION. | | | | | |
| CC | -1- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. | | | | | |
| CC | -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. | | | | | |
| CC | -1- PPM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE | | | | | |
| CC | BONDS. | | | | | |
| CC | -1- SIMILARITY: WITH OTHER BETA CHAINS FROM THE INTEGRIN FAMILY OF | | | | | |
| CC | CELL SURFACE RECEPTOR. | | | | | |
| DR | EMBL; M35197; G553845; -. | | | | | |
| DR | EMBL; A26611; E199079; -. | | | | | |
| DR | PIR; B37057; B37057. | | | | | |
| DR | HSSP; P04355; 2MRT. | | | | | |
| DR | PROSITE; PS00243; INTEGRIN_BETA. 2. | | | | | |
| DR | PROSITE; PS00022; EGF_L; UNKNOWN. 2. | | | | | |
| DR | PROSITE; PS01186; EGF_2; UNKNOWN. 1. | | | | | |
| RW | INTEGRIN; CELL ADHESION; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; | | | | | |
| RW | EXTRACELLULAR MATRIX; CYTOSKELETON. | | | | | |
| FT | NON_TER | 1 | | | | |
| FT | DOMAIN | <1 | 566 | | | |
| FT | DOMAIN | 567 | >577 | | | |
| FT | DOMAIN | 315 | 478 | | | |
| FT | REPEAT | 315 | 360 | | | |
| FT | REPEAT | 361 | 402 | | | |
| FT | REPEAT | 403 | 441 | | | |
| FT | REPEAT | 442 | 478 | | | |
| FT | CARBOHYD | 119 | 119 | | | |
| FT | CARBOHYD | 246 | 246 | | | |
| FT | CARBOHYD | 255 | 255 | | | |
| FT | | | | | | |

| | | | | |
|-------------|------------------|---------|---------------------|-----------------------|
| FT | CARBOHYD | 277 | 277 | POTENTIAL. |
| FT | CARBOHYD | 322 | 322 | POTENTIAL. |
| FT | CARBOHYD | 330 | 330 | POTENTIAL. |
| FT | CARBOHYD | 400 | 400 | POTENTIAL. |
| FT | CARBOHYD | 434 | 434 | POTENTIAL. |
| FT | CARBOHYD | 555 | 555 | POTENTIAL. |
| FT | CARBOHYD | 577 | 577 | POTENTIAL. |
| NON_TER | | 577 | | |
| SQ | SEQUENCE | 577 AA; | 62298 MW; | AEl02507 CRC32; |
| Query Match | | 5.6%; | Score 97; | DB.1; Length 577; |
| Best Local | Similarity | 39.5%; | Pred. No. 2.74e-02; | |
| Matches | 15; Conservative | | 8; Mismatches | 11; Indels 4; Gaps 4; |

Db 398 CDNFSCVRHKGLLCGDNGDC-ECGEYCRSGWTGEYCN 434
| : | | : : : | | : : | | : | : | :
QY 193 CAKF-C-RPRDDSEF-GHSTGSETGEIICLTGNGGDYCH 227

| RESULT | 11 | STANDARD; | PRT; | 3672 AA. |
|--------|---------------------------------------------------------------------|-----------|------|------------------------------------|
| AC | LM2L CAEEL | | | |
| DT | 021313; | | | |
| DT | 01-NOV-1997 (REL. 35, CREATED) | | | |
| DT | 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE) | | | |
| DT | 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE) | | | |
| DE | LAMININ-LIKE PROTEIN K08C7.3 PRECURSOR. | | | |
| GN | K08C7.3. | | | |
| OS | CAENORHABDITIS ELEGANS. | | | |
| OC | EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA. | | | |
| RN | (1) | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-BRISTOL N2; | | | |
| RA | BEKRS M.; | | | |
| RL | SUBMITTED (MAR-1996) TO EMBL/GENBANK/DDJ DATA BANKS. | | | |
| CC | -1- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI). | | | |
| CC | -1- SIMILARITY: CONTAINS ONE LAMININ EGF-LIKE DOMAINS. | | | |
| CC | -1- SIMILARITY: CONTAINS ONE LAMININ DOMAIN IV. | | | |
| DR | EMBL; Z70286; E233841; - | | | |
| DR | WORMPEP; K08C7.3; CE06136. | | | |
| DR | PROSITE; PS00023; EGF_1; 19. | | | |
| DR | PROSITE; PS01186; EGF_2; 4. | | | |
| KM | PROSITE; PS01248; LAMININ_TYPE_EGF; 21. | | | |
| DR | HYPOTHETICAL PROTEIN; LAMININ_EGF-LIKE DOMAIN; SIGNAL; REPEAT. | | | |
| FT | SIGNAL | 1 | 27 | POTENTIAL. |
| FT | CHAIN | 28 | 3672 | LAMININ-LIKE PROTEIN K08C7.3. |
| FT | DOMAIN | 28 | 297 | LAMININ N-TERMINAL (DOMAIN VI). |
| FT | DOMAIN | 298 | 356 | LAMININ EGF-LIKE 1. |
| FT | DOMAIN | 357 | 426 | LAMININ EGF-LIKE 2. |
| FT | DOMAIN | 427 | 471 | LAMININ EGF-LIKE 3. |
| FT | DOMAIN | 472 | 518 | LAMININ EGF-LIKE 4. |
| FT | DOMAIN | 519 | 563 | LAMININ EGF-LIKE 5. |
| FT | DOMAIN | 564 | 609 | LAMININ EGF-LIKE 6. |
| FT | DOMAIN | 610 | 655 | LAMININ EGF-LIKE 7. |
| FT | DOMAIN | 656 | 700 | LAMININ EGF-LIKE 8. |
| FT | DOMAIN | 701 | 755 | LAMININ EGF-LIKE 9. |
| FT | DOMAIN | 756 | 808 | LAMININ EGF-LIKE 10. |
| FT | DOMAIN | 809 | 839 | LAMININ EGF-LIKE 11. (INCOMPLETE). |
| FT | DOMAIN | 1415 | 1460 | LAMININ EGF-LIKE 12. |
| FT | DOMAIN | 1461 | 1505 | LAMININ EGF-LIKE 13. |
| FT | DOMAIN | 1506 | 1553 | LAMININ EGF-LIKE 14. |
| FT | DOMAIN | 1554 | 1604 | LAMININ EGF-LIKE 15. |
| FT | DOMAIN | 1605 | 1614 | LAMININ EGF-LIKE 16. |
| FT | DOMAIN | 1615 | 1796 | LAMININ EGF-LIKE 17. |
| FT | DOMAIN | 1797 | 1829 | LAMININ EGF-LIKE 18. |
| FT | DOMAIN | 1830 | 1879 | LAMININ EGF-LIKE 19. |
| FT | DOMAIN | 1880 | 1936 | LAMININ EGF-LIKE 20. |
| FT | DOMAIN | 1937 | 1989 | LAMININ EGF-LIKE 21. |
| FT | DOMAIN | 1990 | 2036 | LAMININ EGF-LIKE 22. |
| FT | DOMAIN | 2037 | 2083 | LAMININ EGF-LIKE 23. |
| FT | DOMAIN | 2084 | 2131 | LAMININ EGF-LIKE 24. |
| FT | DISULFID | 298 | 307 | BY SIMILARITY. |
| FT | DISULFID | 300 | 320 | BY SIMILARITY. |
| FT | DISULFID | 322 | 331 | BY SIMILARITY. |
| FT | DISULFID | 334 | 354 | BY SIMILARITY. |

CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
DR EMBL: 212171; G50717; -
DR EMBL: U15980; G562108; -
DR EMBL: L12721; G309093; -
DR EMBL: S71340; E120487; -
DR MGD; MGI:94900; DKL1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00022; EGF_1; 5.
DR PROSITE: PS01186; EGF_2; 6.
KM ANTIGEN: TRANSMEMBRANE, EGF-LIKE DOMAIN, REPEAT, GLYCOPROTEIN, SIGNAL.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 385 DELTA-LIKE PROTEIN.
FT PEPTIDE 24 305 FETAL ANTIGEN 1.
FT DOMAIN 24 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 306 329 POTENTIAL.
FT DOMAIN 330 385 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 22 55 EGF-LIKE 1.
FT DOMAIN 33 86 EGF-LIKE 2.
FT DOMAIN 88 125 EGF-LIKE 3.
FT DOMAIN 127 168 EGF-LIKE 4.
FT DOMAIN 172 208 EGF-LIKE 5.
FT DOMAIN 210 247 EGF-LIKE 6.
FT DISULFID 26 37 BY SIMILARITY.
FT DISULFID 30 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 57 68 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 76 85 BY SIMILARITY.
FT DISULFID 92 103 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 156 BY SIMILARITY.
FT DISULFID 158 167 BY SIMILARITY.
FT DISULFID 176 187 BY SIMILARITY.
FT DISULFID 181 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 214 225 BY SIMILARITY.
FT DISULFID 219 235 BY SIMILARITY.
FT DISULFID 237 246 BY SIMILARITY.
FT CARBOHYD 100 100 POTENTIAL.
FT CARBOHYD 295 295 POTENTIAL.
FT CARBOHYD 342 342 POTENTIAL.
FT CONFLICT 250 250 R -> P (IN REF. 2).
FT CONFLICT 320 385 VIGTVAIVELINKCTEWSNLRYNHFRRKKMLLOYNGEE
LAVNIIFPEKIDMTFNKREADEEL -> CHAPPSSSTIS
AKRCPGTCATVTCFARRRTSCSITAAWSRISSSPRRLT
(IN REF. 2).
SQ SEQUENCE 385 AA; 41320 MW; 8630456D CRC32;
Query Match 5.4%; Score 95; DB 1; length 385;
Best Local Similarity 30.5%; Pred. No. 5.81e-02;
Matches 18; Conservative 9; Mismatches 30; Indels 2; Gaps 2;
Db 31 DPQGFCEADWVCRCHVWGEGPLCDK-CVTAPGCV-NGCVCKPMQICKDGKGFCEI 87
OY 170 ESQYTSLEFDPRVYCDLNYTSGCAKFCRPRDDSGHSTCSTGIIILITGQSDYCHI 228
RESULT 13
ID VL25 CAEEL STANDARD; PRT; 810 AA.
AC P34418:
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 89.2 KD PROTEIN F42H10.5 IN CHROMOSOME III.
GN F42H10.5
OS CAENORHADITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEA; RHADITIDA.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,
RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,
RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,
RA SUTSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,
RA WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,
RA WOHLDMAN P.;
RL NATURE 368:32-38(1994).
DR EMBL: L08403; G289681; -
DR WORKPEP: F42H10.5; CE00167.
KM HYPOTHETICAL PROTEIN.
SQ SEQUENCE 810 AA; 89154 MW; 611370C4 CRC32;
Query Match 5.4%; Score 94; DB 1; length 810;
Best Local Similarity 23.4%; Pred. No. 8.43e-02;
Matches 15; Conservative 16; Mismatches 31; Indels 2; Gaps 2;
Db 676 GGCEALAIYGDLPQISNGSESK-ENIVNGYFDEISSTVESWFMRTGPNQAPL 734
OY 59 GSKTRFRCCLKHYATDTTSQCTYGVDTPIIGENSVNLTDARFQNKFTNPIQFF 118
Db 735 SY-W 737
OY 119 SFSW 122
RESULT 14
ID PAC4 HUMAN STANDARD; PRT; 969 AA.
AC P29122:
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE SUBTILISIN-LIKE PROTEASE PACE4 PRECURSOR (EC 3.4.21.-).
OS PACE4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; PRIMATES.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE: 92075167.
RA KIEFER M.C., TUCKER J.E., JOH R., LANDSBERG K.E., SALTMAN D.,
RA BARR P.J.;
RL DNA CELL BIOL. 10:757-769(1991).
CC CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: WIDESPREAD, WITH COMPARATIVELY HIGHER LEVELS
CC IN THE LIVER. PACE4.1 WAS ONLY FOUND IN THE LIVER.
CC -1- ALTERNATIVE PRODUCTS: TWO VARIANTS SEEM TO BE PRODUCED BY
CC ALTERNATIVE SPLICING OF A SINGLE GENE. THEY DIFFER AFTER
CC LVS-471.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILISIN FAMILY. HIGH SIMILARITY WITH OTHER FURIN-LIKE ENZYMES.
DR EMBL: M80482; G189532; -
DR PIR: A39490; A39490.
DR HSSP: 099405; 1MPT.
DR MIM: 167405; -
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KM HYDROLASE, SERINE PROTEASE; GLYCOPROTEIN; ZIMOGEN; SIGNAL;
KM ALTERNATIVE SPLICING; REPEAT.
FT SIGNAL 1 63 POTENTIAL.
FT PROPEP 64 149 SUBTILISIN-LIKE PROTEASE PACE4.
FT CHAIN 150 969 CYS-RICH REGION.
FT DOMAIN 695 969 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 205 205 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 246 246

FT ACT_SITE 420 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 553 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 686 941 5 X TANDEM REPEATS, CYS-RICH.
FT REPEAT 686 739 1.
FT REPEAT 740 790 2.
FT REPEAT 791 838 3.
FT REPEAT 839 887 4.
FT REPEAT 888 941 5.
FT CARBOHYD 259 POTENTIAL.
FT CARBOHYD 914 914 POTENTIAL.
FT CARBOHYD 932 POTENTIAL.
FT VARSPLIC 471 471 K -> KGAFAFWTWIGWPNV (IN PACE4.1).
FT VARSPLIC 472 969 MISSING (IN PACE4.1).
SQ SEQUENCE 969 AA; 106419 MW; 1FE74A29 CRC32;

Query Match 5.4%; Score 94; DB 1; Length 969;
Best Local Similarity 38.6%; Pred. No. 8.43e-02;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;

Db 733 CPLGYFGDTAARRCR-RCHK-GCETCSSRAATQCLSCRRGFYHH 774

Qy 184 CDLNYGSCAKFCRPRDDSFHSGTCTGETGIIICLTGWGQDYCH 227

RESULT 15
ID CANL_YEAST STANDARD; PRT; 590 AA.
AC P04817;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ARGinine PERMEASE.
GN CANL OR YEL063C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88165106.
RA AHMAD M., BUSSEY H.;
RL CURR. GENET. 10:587-592(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 8608235.
RA HOFEMAN W.;
RL J. BIOL. CHEM. 260:11831-11837(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DIETRICH F. S., MULLIGAN J. T., HENNESSEY K. M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J. M.,
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICK-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,
RA PETEL F. X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R. W.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -!- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGinine.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES.
DR EMBL: X03784; G3442; -.
DR EMBL: M11724; G171155; -.
DR EMBL: U18795; G603255; -.
DR PIR: A23922; ORBYPR.
DR SGD: L0000213; CAN1.
DR PROSITE: PS00218; AMINO-ACID TRANSPORT: TRANSMEMBRANE.
KW TRANSPORT; AMINO-ACID TRANSPORT: TRANSMEMBRANE.
FT DOMAIN 1 92 HYDROPHILIC.
FT TRANSMEM 93 110 POTENTIAL.
FT TRANSMEM 230 254 POTENTIAL.
FT TRANSMEM 324 343 POTENTIAL.
FT TRANSMEM 377 396 POTENTIAL.
FT TRANSMEM 420 437 POTENTIAL.
FT TRANSMEM 451 469 POTENTIAL.

FT TRANSMEM 494 518 POTENTIAL.
FT TRANSMEM 525 548 POTENTIAL.
FT DOMAIN 549 590 HYDROPHILIC.
FT CONFLICT 534 534 I -> V (IN REF. 1).
SQ SEQUENCE 590 AA; 65785 MW; 74DCD8FC CRC32;

Query Match 5.3%; Score 92; DB 1; Length 590;
Best Local Similarity 37.0%; Pred. No. 1.76e-01;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 157 SQRLSPAFGAANGYMYWFSWATFAL 183
Qy 102 AQRFQNKGFNPQPFPSFSGMPTFSL 128

Search completed: Fri Jun 12 12:05:13 1998
Job time : 20 secs.

FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 832 DELTA PROTEIN.
FT DOMAIN 26 595 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 569 617 POTENTIAL.
FT DOMAIN 618 832 INTRACELLULAR (POTENTIAL).
FT DOMAIN 217 566 9 EGF-TYPE REPEATS.
FT REPEAT 217 256 EGF-LIKE 1.
FT REPEAT 257 290 EGF-LIKE 2.
FT REPEAT 291 330 EGF-LIKE 3.
FT REPEAT 331 373 EGF-LIKE 4.
FT REPEAT 374 417 EGF-LIKE 5.
FT REPEAT 418 452 EGF-LIKE 6.
FT REPEAT 453 490 EGF-LIKE 7.
FT REPEAT 491 528 EGF-LIKE 8.
FT REPEAT 529 566 EGF-LIKE 9.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 137 137 POTENTIAL.
FT CARBOHYD 167 167 POTENTIAL.
FT CARBOHYD 421 421 POTENTIAL.
FT CARBOHYD 649 649 POTENTIAL.
FT CONFLICT 437 438 GK -> ET (IN REF. 2).
FT CONFLICT 443 443 A -> S (IN REF. 2).
FT CONFLICT 459 459 G -> A (IN REF. 2).
FT CONFLICT 490 490 S -> T (IN REF. 2).
FT CONFLICT 591 591 T -> A (IN REF. 2).
SQ SEQUENCE 832 AA; 88813 MW; CF9ABEC1 CRC32;
Query Match 100.0%; Score 1734; DB 3; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWIKLLTAFCTFVIVQVHSGSFELRLKYFSDHGRNDRCCSGSDGATGKCLGS 60
Qy 1 MHWIKLLTAFCTFVIVQVHSGSFELRLKYFSDHGRNDRCCSGSDGATGKCLGS 60
Db 61 CKTRFVCLKHQYATIDTTSQCTGYGVITPILGENSVNLTDQRFQNGKFTNPIQPFPSF 120
Qy 61 CKTRFVCLKHQYATIDTTSQCTGYGVITPILGENSVNLTDQRFQNGKFTNPIQPFPSF 120
Db 121 SWPGFTSLIVEAWHDTNNSGNARTNKLIIORLLVQVLEVSSEWTKNKSQYTSLEYDF 180
Qy 121 SWPGFTSLIVEAWHDTNNSGNARTNKLIIORLLVQVLEVSSEWTKNKSQYTSLEYDF 180
Db 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPK 230
Qy 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPK 230

RESULT 2
ID Q90656 PRELIMINARY; PRT; 728 AA.
AC Q90656;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE PROTEIN C-DELTA-1.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE-SPINAL CORD;
RX MEDLINE; 95319507.
RA HENRIQUE D., ADAM J., WYAT A., CHITNIS A., LEWIS J., ISH-HOROWICZ D.;
RL NATURE 375:787-790(1995).
DR EMBL; U26590; G882412; -.
DR PROSITE; PS01187; EGF CA: 2.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 728 AA; 79861 MW; 7439F575 CRC32;

Query Match 44.7%; Score 775; DB 12; Length 728;
Best Local Similarity 49.1%; Pred. No. 1.54e-162;
Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1;

Db 20 QYDGGVFEKLQEFVNNKGLLSNRNCCRGPGGAGQOOCCKTFRRVCLKHYQASVSP 79
Qy 19 QVHSSGSFELRLKYFSDHGRNDRCCSGSDGATGKCLGSCKTRFRVCLKHYQATIDT 78
Db 80 EPPCTYSATIPVLGANSFSPDAGGADPASNIRFPFGFTWPTGTFSLIIIEALHTDSP 139
Qy 79 TSQCTYGDVITPILGENSVNLTDQRFQNGKFTNPIQPFPSFWPTGTFSLIVEAWH-DTN 137
Db 140 DDLTTENPERLISRLATQRHLAVGEWSQDLHSSGRTDLKYSYRFVCDHYHGGCVFPC 199
Qy 138 NSGNARTNKLIIORLLVQVLEVSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFC 197
Db 200 RPDRDRFGHFTGCGEKVCNPGWKGQYCTEP 231
Qy 198 RPDRDSFGHSTCSETGEIICLTGWQGDYCHIP 229

RESULT 3
ID P87357 PRELIMINARY; PRT; 717 AA.
AC P87357;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DELTAD TRANSMEMBRANE PROTEIN PRECURSOR.
GN DELTAD.
OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RA DORSEIFER P., TAKKE C., CAMPOS-ORTEGA J.A.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y11760; E307461; -.
DR PROSITE; PS01187; EGF CA: 2.
KW SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 717 DELTAD TRANSMEMBRANE PROTEIN.
SQ SEQUENCE 717 AA; 79061 MW; 5CC32ECA CRC32;

Query Match 44.3%; Score 769; DB 12; Length 717;
Best Local Similarity 46.0%; Pred. No. 5.29e-161;
Matches 103; Conservative 40; Mismatches 74; Indels 7; Gaps 5;

Db 5 MIAYLVC-VMSQFCGSGVFEKLQEFVNNKGLLSNRNCCRGPGGAGQOOCCKTFRRVCLKHYQASVSP 79
Qy 7 LTAFCTFVIVQVHSGSFELRLKYFSDHGRNDRCCSGSDGATGKCLGSCKTRFR 66
Db 61 ICLKHQYQANVSPDPCCTGYGAVTPVLGNSNFQVPS--FPDSSFTNPFPFAFGFTWPTGTF 118
Qy 67 VCLKHQYATIDTTSQCTGYGVITPILGENSVNLTDQRFQNGKFTNPIQPFPSFWPTGTF 126
Db 119 SLIIIEALHTDSTDDLSTENPDRLISRTQRHLTVGEWSQDLQVGGRTKLYSYRFVCD 178
Qy 127 SLIVEAWH-DTNNSGNARTNKLIIORLLVQVLEVSSEWTKNKSQYTSLEYDFRVTCD 185
Db 179 EHYHGGCVFPCPRDDTFGHFTCGERGEIICNCSGWKGQYCTEP 222
Qy 186 LNYGSGCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIP 229

RESULT 4
ID Q00548 PRELIMINARY; PRT; 723 AA.
AC Q00548;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DELTA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MANN R.S., GRAY G.E., HENRIQUE D., ISH-HOROWICZ D.,

| MPsrch_nn | n.a. | - | n.a. | database search, using Smith-Waterman algorithm |
|-----------|------|----|------|-------------------------------------------------|
| 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 2 | 2 |
| 3 | 3 | 3 | 3 | 3 |
| 4 | 4 | 4 | 4 | 4 |
| 5 | 5 | 5 | 5 | 5 |
| 6 | 6 | 6 | 6 | 6 |
| 7 | 7 | 7 | 7 | 7 |
| 8 | 8 | 8 | 8 | 8 |
| 9 | 9 | 9 | 9 | 9 |
| 10 | 10 | 10 | 10 | 10 |
| 11 | 11 | 11 | 11 | 11 |
| 12 | 12 | 12 | 12 | 12 |
| 13 | 13 | 13 | 13 | 13 |
| 14 | 14 | 14 | 14 | 14 |
| 15 | 15 | 15 | 15 | 15 |
| 16 | 16 | 16 | 16 | 16 |
| 17 | 17 | 17 | 17 | 17 |
| 18 | 18 | 18 | 18 | 18 |
| 19 | 19 | 19 | 19 | 19 |
| 20 | 20 | 20 | 20 | 20 |
| 21 | 21 | 21 | 21 | 21 |
| 22 | 22 | 22 | 22 | 22 |
| 23 | 23 | 23 | 23 | 23 |
| 24 | 24 | 24 | 24 | 24 |
| 25 | 25 | 25 | 25 | 25 |
| 26 | 26 | 26 | 26 | 26 |
| 27 | 27 | 27 | 27 | 27 |
| 28 | 28 | 28 | 28 | 28 |
| 29 | 29 | 29 | 29 | 29 |
| 30 | 30 | 30 | 30 | 30 |
| 31 | 31 | 31 | 31 | 31 |
| 32 | 32 | 32 | 32 | 32 |
| 33 | 33 | 33 | 33 | 33 |
| 34 | 34 | 34 | 34 | 34 |
| 35 | 35 | 35 | 35 | 35 |
| 36 | 36 | 36 | 36 | 36 |
| 37 | 37 | 37 | 37 | 37 |
| 38 | 38 | 38 | 38 | 38 |
| 39 | 39 | 39 | 39 | 39 |
| 40 | 40 | 40 | 40 | 40 |
| 41 | 41 | 41 | 41 | 41 |
| 42 | 42 | 42 | 42 | 42 |
| 43 | 43 | 43 | 43 | 43 |
| 44 | 44 | 44 | 44 | 44 |
| 45 | 45 | 45 | 45 | 45 |
| 46 | 46 | 46 | 46 | 46 |
| 47 | 47 | 47 | 47 | 47 |
| 48 | 48 | 48 | 48 | 48 |
| 49 | 49 | 49 | 49 | 49 |
| 50 | 50 | 50 | 50 | 50 |
| 51 | 51 | 51 | 51 | 51 |
| 52 | 52 | 52 | 52 | 52 |
| 53 | 53 | 53 | 53 | 53 |
| 54 | 54 | 54 | 54 | 54 |
| 55 | 55 | 55 | 55 | 55 |
| 56 | 56 | 56 | 56 | 56 |
| 57 | 57 | 57 | 57 | 57 |
| 58 | 58 | 58 | 58 | 58 |
| 59 | 59 | 59 | 59 | 59 |
| 60 | 60 | 60 | 60 | 60 |
| 61 | 61 | 61 | 61 | 61 |
| 62 | 62 | 62 | 62 | 62 |
| 63 | 63 | 63 | 63 | 63 |
| 64 | 64 | 64 | 64 | 64 |
| 65 | 65 | 65 | 65 | 65 |
| 66 | 66 | 66 | 66 | 66 |
| 67 | 67 | 67 | 67 | 67 |
| 68 | 68 | 68 | 68 | 68 |
| 69 | 69 | 69 | 69 | 69 |
| 70 | 70 | 70 | 70 | 70 |
| 71 | 71 | 71 | 71 | 71 |
| 72 | 72 | 72 | 72 | 72 |
| 73 | 73 | 73 | 73 | 73 |
| 74 | 74 | 74 | 74 | 74 |
| 75 | 75 | 75 | 75 | 75 |
| 76 | 76 | 76 | 76 | 76 |
| 77 | 77 | 77 | 77 | 77 |
| 78 | 78 | 78 | 78 | 78 |
| 79 | 79 | 79 | 79 | 79 |
| 80 | 80 | 80 | 80 | 80 |
| 81 | 81 | 81 | 81 | 81 |
| 82 | 82 | 82 | 82 | 82 |
| 83 | 83 | 83 | 83 | 83 |
| 84 | 84 | 84 | 84 | 84 |
| 85 | 85 | 85 | 85 | 85 |
| 86 | 86 | 86 | 86 | 86 |
| 87 | 87 | 87 | 87 | |

```
Run on: Sat Jun 13 06:53:25 1998; MASPAR time 3140.54 Seconds
1536 637 williamson updates/sec
```

Tabular output not generated.

| | |
|----------------|---------------------------------------------------|
| Title: | >US-08-083-590A-1 |
| Description: | (1-2892) from US08083590A.seq |
| Perfect Score: | 2892 |
| N.A. Sequence: | 1 GAATTCGAGGAAATTATCA.....GCGTTTCGATCCGAATTC 2892 |
| Comp: | CTTAAGCCTCTATAAGT.....CGGCAACGTACGCGTTAAG |

Scoring table: TABLE default
Gap 6

```
Nmatch  STD :  Dbase 0;  Query 0
```

Searched: 457396 seqs, 834342348 bases x 2

```
Post-processing: Minimum Match 08
                  Listing first 45 summaries
```

```
Database: embi54
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro 12:em_vl
genbank106
```

Statistics: Mean 11.940; Variance 6.492; scale 1.8339

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

| Result | Query |
|--------|-------|
| ... | ... |

| No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|-----|-------|-------|--------|----|-----------|-------------------------|-----------|
| 1 | 2892 | 100.0 | 2892 | 15 | DROEGFLUA | Drosophila sp. epiderm | 0.00e+00 |
| 2 | 2892 | 100.0 | 2892 | 18 | I56055 | Sequence 5 from patient | 0.00e+00 |
| 3 | 2871 | 99.3 | 2889 | 15 | DMEDELTA | D.melanogaster D1 mRNA | 0.00e+00 |
| 4 | 2746 | 99.0 | 4724 | 15 | DMDLGH | Drosophila mRNA for De | 0.00e+00 |
| 5 | 564 | 19.5 | 600 | 15 | DMDLEGN | Drosophila melanogaster | 0.00e+00 |
| 6 | 169 | 5.8 | 2187 | 17 | GCU26550 | Gallus gallus C-Delta- | 3.91e-97 |
| 7 | 131 | 4.5 | 3162 | 22 | AF003522 | Homio sapiens Delta mRN | 4.31e-66 |
| 8 | 115 | 4.0 | 2342 | 17 | DREDETRFD | D. rerio mRNA for Delta | 1.72e-57 |
| 9 | 111 | 3.8 | 2795 | 23 | RND7R8889 | Rattus norvegicus Delt | 1.25e-54 |
| 10 | 108 | 3.7 | 2857 | 23 | KMDDELTA1 | M.musculus mRNA for De | 1.81e-52 |
| 11 | 100 | 3.5 | 3007 | 17 | AF006488 | Danio rerio delta mRNA | 8.92e-42 |
| 12 | 94 | 3.3 | 2166 | 17 | XELXDELTA | Xenopus laevis X-Delta | 1.53e-42 |
| 13 | 79 | 2.7 | 3051 | 17 | AF030031 | Danio rerio Delta (de | 3.92e-32 |
| 14 | 71 | 2.5 | 7218 | 18 | I66494 | Sequence 14 from patien | 1.05e-22 |
| 15 | 70 | 2.4 | 2869 | 17 | XLU70843 | Xenopus laevis Notch 1 | 4.90e-26 |

| | | | | | | | |
|----|----|-----|-------|----|------------|-------------------------|----------|
| 16 | 63 | 2.2 | 3362 | 15 | SUSEGFI | Strongylocentrotus pur | 2.15e-21 |
| 17 | 62 | 2.1 | 7646 | 15 | AFPO0664 | Lytechinus variegatus | 9.65e-21 |
| 18 | 57 | 2.0 | 1483 | 15 | SUSGEF | S.purpuratus epidermal | 1.67e-17 |
| 19 | 57 | 2.0 | 1580 | 15 | SUSEGFI.I | Strongylocentrotus pur | 1.67e-17 |
| 20 | 57 | 2.0 | 3609 | 23 | MMNOTB | M.musculus notch B mRNA | 1.67e-17 |
| 21 | 57 | 2.0 | 5575 | 23 | RATJARG | Rattus norvegicus jag | 1.67e-17 |
| 22 | 57 | 2.0 | 7332 | 21 | HUMTANI | Human TNF-1 mRNA (homo | 1.67e-17 |
| 23 | 57 | 2.0 | 7579 | 23 | D32210 | Mus musculus (Notch2) | 1.67e-17 |
| 24 | 58 | 2.0 | 10452 | 15 | DRONOT05 | D.melanogaster growth | 3.81e-18 |
| 25 | 54 | 1.9 | 1574 | 15 | DROSLITEGF | Drosophila melanogaste | 1.36e-15 |
| 26 | 56 | 1.9 | 5274 | 15 | DMSER | D.melanogaster serrate | 7.29e-17 |
| 27 | 54 | 1.9 | 5401 | 15 | DMSLIT | Drosophila mRNA for sl | 1.36e-15 |
| 28 | 54 | 1.9 | 5545 | 15 | DROTMPLA | D.melanogaster serrate | 1.36e-15 |
| 29 | 54 | 1.9 | 8287 | 23 | RATNOTCHX | Rat notch 2 mRNA. | 1.36e-15 |
| 30 | 56 | 1.9 | 17137 | 15 | DRONOTCH03 | D.melanogaster Notch 1 | 7.29e-17 |
| 31 | 51 | 1.8 | 2101 | 15 | SUEFIBIT1 | Sea urchin fibropellin | 1.04e-13 |
| 32 | 52 | 1.8 | 2388 | 15 | SUEFIBIT1 | Sea urchin fibropellin | 2.46e-14 |
| 33 | 51 | 1.8 | 4635 | 22 | HSY14330 | Homo sapiens partial m | 1.04e-13 |
| 34 | 51 | 1.8 | 4702 | 22 | AFPO03521 | Homo sapiens Jagged 2 | 1.04e-13 |
| 35 | 51 | 1.8 | 4749 | 22 | AFD20201 | Homo sapiens Jagged 2 | 1.04e-13 |
| 36 | 53 | 1.8 | 4855 | 22 | HS061276 | Human transmembrane pr | 5.80e-15 |
| 37 | 53 | 1.8 | 4860 | 22 | AFD29777 | Homo sapiens hJAG2.del | 1.04e-13 |
| 38 | 51 | 1.8 | 4974 | 22 | AFD29778 | Homo sapiens jagged2 (| 1.04e-13 |
| 39 | 51 | 1.8 | 5445 | 22 | HS0777720 | Human transmembrane pr | 1.04e-13 |
| 40 | 51 | 1.8 | 5457 | 22 | AFD28593 | Homo sapiens transmemb | 1.04e-13 |
| 41 | 51 | 1.8 | 5590 | 22 | HS073936 | Homo sapiens Jagged 1 | 1.04e-13 |
| 42 | 51 | 1.8 | 5942 | 22 | AFD03837 | Homo sapiens Jagged1 (| 1.04e-13 |
| 43 | 50 | 1.7 | 3582 | 17 | GCGSERAT | G.gallus mRNA for C-se | 4.33e-13 |
| 44 | 50 | 1.7 | 7471 | 17 | BRNOTCH | B.perio Notch mRNA. | 4.33e-13 |
| 45 | 48 | 1.7 | 2551 | 17 | AB004629 | Fugu rubripes gene for | 7.37e-12 |

ALIGNMENTS

| | | | | | | |
|------------|---|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|-----|-----|-------------|
| RESULT | 1 | DROEGFLEA | 2892 bp | DNA | INV | 22-JUN-1994 |
| LOCUS | | Drosophila sp. | epidermal growth factor-like element (Notch and Delta) gene regions. | | | |
| DEFINITION | | L03725 | | | | |
| KEYWORDS | | g290222 | | | | |
| ACCESSION | | Delta gene; Notch gene; cell adhesion activity; epidermal growth factor-like protein. | | | | |
| NID | | Drosophila sp. DNA. | | | | |
| SOURCE | | Drosophila sp. | | | | |
| ORGANISM | | Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda; Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. | | | | |
| REFERENCE | | 1 (bases 1 to 2892) | | | | |
| AUTHORS | | Lieber,T., Wesley,C.S., Alcamo,E., Hassel,B., Krane,J.F., Campos-Orenga,J.A. and Young,M.W. | | | | |
| TITLE | | Single amino acid substitutions in EGF-like elements of Notch and Delta modify Drosophila development and affect cell adhesion in vitro | | | | |
| JOURNAL | | Neuron 9 (5), 847-859 (1992) | | | | |
| MEDLINE | | 93040226 | | | | |
| FEATURES | | Location/Qualifiers | | | | |
| Source | | 1..2892 | | | | |
| | | /organism="Drosophila sp." | | | | |
| | | /db_xref="taxon:7242" | | | | |
| | | 1..2892 | | | | |
| gene | | /note="Delta" | | | | |
| | | /gene="DL" | | | | |
| | | /allele="5" | | | | |
| | | /db_xref="FlyBase:FBNgn000463" | | | | |
| | | replace(1054,"") | | | | |
| mutation | | /gene="DL" | | | | |
| | | /note="a' in mutation" | | | | |
| | | /label-DlSup5 | | | | |
| | | /db_xref="FlyBase:FBNgn000463" | | | | |
| | | replace(1699,"") | | | | |
| mutation | | /gene="DL" | | | | |
| | | /note="a' in mutation" | | | | |
| | | /label-DlSup4 | | | | |

variation /db_xref="FlyBase:FBgn0000463"
replace(2096,"")
/gene="Dl"
/note="amino acid G to A in Dlsup4, Dlsup5 and one other strain; 'c' in variation"
/db_xref="FlyBase:FBgn0000463"
replace(2125,"")
/gene="Dl"
/note="amino acid L to M in a strain; 'a' in variation"
/db_xref="FlyBase:FBgn0000463"
replace(2576,"")
/gene="Dl"
/note="amino acid A to V in 4 strains; 't' in variation"
/db_xref="FlyBase:FBgn0000463"
739 a 764 c 808 g 581 t

BASE COUNT 739 a 764 c 808 g 581 t
ORIGIN
Query Match 100.0%; Score 2892; DB 15; Length 2892;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 GAATTCGGAGGAATTATTCAAACATATAACACATAATAAATTTGAGTAGTCCGCACA 60
Qy 1 GAATTCGGAGGAATTATTCAAACATATAACACATAATAAATTTGAGTAGTCCGCACA 60
Db 61 CACACACACACAGCCCGTGGATTATTACACTAAAGCGACACTCAATCCAAAATCA 120
Qy 61 CACACACACACAGCCCGTGGATTATTACACTAAAGCGACACTCAATCCAAAATCA 120
Db 121 GCAACAAAACATCAATAAACATGCAATGGATTAAATGTTTATTAAACAGATTCAATTCG 180
Qy 121 GCAACAAAACATCAATAAACATGCAATGGATTAAATGTTTATTAAACAGATTCAATTCG 180
Db 181 TTCACAGTCATCGTGAGGTTTACAGTCCGGCAGCTTTGAGTTGGCCCTGAAAGTACTTC 240
Qy 181 TTCACAGTCATCGTGAGGTTTACAGTCCGGCAGCTTTGAGTTGGCCCTGAAAGTACTTC 240
Db 241 AGCAACGATCAGGGCGGCGGACACAGAGGTCGCTGTCAGCGGGGAGTCGGACGAGCG 300
Qy 241 AGCAACGATCAGGGCGGCGGACACAGAGGTCGCTGTCAGCGGGGAGTCGGACGAGCG 300
Db 301 ACGGCAAGTCGTCGGCAGCTGCAAGACGCGGTTTCGCTCTGCTTAAAGCACTACCAG 360
Qy 301 ACGGCAAGTCGTCGGCAGCTGCAAGACGCGGTTTCGCTCTGCTTAAAGCACTACCAG 360
Db 361 GCCACCATCGACACCACTCCCGAGTGCACCTACGGGGAGCTGATCAGCCCAATTCGCGC 420
Qy 361 GCCACCATCGACACCACTCCCGAGTGCACCTACGGGGAGCTGATCAGCCCAATTCGCGC 420
Db 421 GAGAACTCGGTCAATCTGACCGCCGAGCGCTCCAGAACAGGGCTTCACGAATCCC 480
Qy 421 GAGAACTCGGTCAATCTGACCGCCGAGCGCTCCAGAACAGGGCTTCACGAATCCC 480
Db 481 ATCCAGTTCCTTCGCTTCATCGCGCGGTACCTTCGCTGATCGTCGAGGCGCTGG 540
Qy 481 ATCCAGTTCCTTCGCTTCATCGCGCGGTACCTTCGCTGATCGTCGAGGCGCTGG 540
Db 541 CATGATACGAACAATAGCGGCAATGCGGAAACCAACAGCTCTCATCCAGCGACTCTTG 600
Qy 541 CATGATACGAACAATAGCGGCAATGCGGAAACCAACAGCTCTCATCCAGCGACTCTTG 600
Db 601 GTGACGAGGTACTGGAGTGTCCTCCGAATGGAAGACGAACAGTCGGAAATCGCAGTAC 660
Qy 601 GTGACGAGGTACTGGAGTGTCCTCCGAATGGAAGACGAACAGTCGGAAATCGCAGTAC 660
Db 661 ACGTCGCTGAGTACGATTTCGCTGTCACCTCGGATCTCAACTACTACGATCCGCTGT 720
Qy 661 ACGTCGCTGAGTACGATTTCGCTGTCACCTCGGATCTCAACTACTACGATCCGCTGT 720
Db 721 GCCAAGTTCTGCGGCGCCCGCGACAGATTCAATTTGGACACTTCGATTCGCGAGACGGC 780
Qy 721 GCCAAGTTCTGCGGCGCCCGCGACAGATTCAATTTGGACACTTCGATTCGCGAGACGGC 780

Db 781 GAAATTATCTGTTTACCGGATGGCAGGCGATTACTGTACATACCCAAATCGCCAAA 840
Qy 781 GAAATTATCTGTTTACCGGATGGCAGGCGATTACTGTACATACCCAAATCGCCAAA 840
Db 841 GGTGTGAACATGGACATTCGCGACAAACCAATCAATGCTTTGCCAATCGGCTGGAAG 900
Qy 841 GGTGTGAACATGGACATTCGCGACAAACCAATCAATGCTTTGCCAATCGGCTGGAAG 900
Db 901 GGAGCCTTGTGCAACAGTGCCTTCTGGACCGAATGATGATTCGCAACCTGCAACAAA 960
Qy 901 GGAGCCTTGTGCAACAGTGCCTTCTGGACCGAATGATGATTCGCAACCTGCAACAAA 960
Db 961 CCCTGGACTTGCATCTGCAACAGGCTTGGGAGGCTTGGGAGGCTTGTACTGCAACAGGATCTGAAC 1020
Qy 961 CCCTGGACTTGCATCTGCAACAGGCTTGGGAGGCTTGGGAGGCTTGTACTGCAACAGGATCTGAAC 1020
Db 1021 TACTGCACCAACACAGACCCCTGCAAGATGCGGGAACCTGCTTCAACACCGCGGAGGA 1080
Qy 1021 TACTGCACCAACACAGACCCCTGCAAGATGCGGGAACCTGCTTCAACACCGCGGAGGA 1080
Db 1081 TTGTACACATGCAAAATGCGCTCCAGGATACAGTGGTGGATTCGCAAAATGAGATCTAC 1140
Qy 1081 TTGTACACATGCAAAATGCGCTCCAGGATACAGTGGTGGATTCGCAAAATGAGATCTAC 1140
Db 1141 TCCTGCGATGCCGATGTCAATCCCTGCCAGATGGTGGTACCTGCGATGAGCGGCAC 1200
Qy 1141 TCCTGCGATGCCGATGTCAATCCCTGCCAGATGGTGGTACCTGCGATGAGCGGCAC 1200
Db 1201 ACAAACACCGGCTTACAAAGTGTCAATTCGCGCAACCGGCTGGAGCGAAAGATGTCGAGGAG 1260
Qy 1201 ACAAACACCGGCTTACAAAGTGTCAATTCGCGCAACCGGCTGGAGCGAAAGATGTCGAGGAG 1260
Db 1261 AAGTGCCTCAGTGTTCGGACAAACCTGTCATCAGGGAATCTGCCCAACGTTGCTCCT 1320
Qy 1261 AAGTGCCTCAGTGTTCGGACAAACCTGTCATCAGGGAATCTGCCCAACGTTGCTCCT 1320
Db 1321 GGCTTGGGAGCAAGGTCAGGCTACAGGCTACCACTGCGAATGTCCTTGCCTACACGCGACCC 1380
Qy 1321 GGCTTGGGAGCAAGGTCAGGCTACAGGCTACCACTGCGAATGTCCTTGCCTACACGCGACCC 1380
Db 1381 AACTGGGATCTCCAGCTGGCAAACTGCAGTCCGAATCCATGCAATAACGGTGAAGCTGT 1440
Qy 1381 AACTGGGATCTCCAGCTGGCAAACTGCAGTCCGAATCCATGCAATAACGGTGAAGCTGT 1440
Db 1441 CAGCGAGCGGAAAGTGTATTTGCCAGCGGATTTTCGGGAACGAGATCGGAGACCAAC 1500
Qy 1441 CAGCGAGCGGAAAGTGTATTTGCCAGCGGATTTTCGGGAACGAGATCGGAGACCAAC 1500
Db 1501 ATTGACGATGCTTGGCCACAGTGGGAGACGAGGACCTGTCATAGATATGGTCAAC 1560
Qy 1501 ATTGACGATGCTTGGCCACAGTGGGAGACGAGGACCTGTCATAGATATGGTCAAC 1560
Db 1561 CAATATCGCTGCCAATGCTTCCCGGTTCCATGGCACCCCACTGTAGTAGCAAGTTGAC 1620
Qy 1561 CAATATCGCTGCCAATGCTTCCCGGTTCCATGGCACCCCACTGTAGTAGCAAGTTGAC 1620
Db 1621 TTGTGCTCATCAGACCGTGTGCAATGAGGAACTGCTTGAATCTCAACAACGATTCAC 1680
Qy 1621 TTGTGCTCATCAGACCGTGTGCAATGAGGAACTGCTTGAATCTCAACAACGATTCAC 1680
Db 1681 CAGTGCACCTGCTGCGGGATTACTGGCAAGGATGCTCTGTGGACATCGATGAGTGC 1740
Qy 1681 CAGTGCACCTGCTGCGGGATTACTGGCAAGGATGCTCTGTGGACATCGATGAGTGC 1740
Db 1741 AGCAGTGGACCCCTGTATTAACCGCGGCACTTGCATGACCCGCTCAATCTGTCGAATGC 1800
Qy 1741 AGCAGTGGACCCCTGTATTAACCGCGGCACTTGCATGACCCGCTCAATCTGTCGAATGC 1800
Db 1801 GTGTGTGCCAATGTTTCAGGGGCAAGCAGTGGCATGAGGAGTCTTACGATTCGGTGACC 1860
Qy 1801 GTGTGTGCCAATGTTTCAGGGGCAAGCAGTGGCATGAGGAGTCTTACGATTCGGTGACC 1860
Db 1861 TTGATGCCCCCAATATGAGGCGACCAACAGCGAGACCGGATGTTTGGACCAATGCC 1920

|||||
QY 1861 TTGATGCGCCACCAATATGAGCGACACACAGGAGAGCGGATGGTTGACCAATGCG 1920
Db 1921 CAGGTAGTCTTAATGCTGTTTTCCTGCTGGATGCTTTGGTGGCGTATTTGGCGG 1980
QY 1921 CAGGTAGTCTTAATGCTGTTTTCCTGCTGGATGCTTTGGTGGCGTATTTGGCGG 1980
Db 1981 TGGCTGGCTCTGATGATGAGCGGAGGTAAGCGTGGCTGAGGAAAGAGACCGCGAG 2040
QY 1981 TGGCTGGCTCTGATGATGAGCGGAGGTAAGCGTGGCTGAGGAAAGAGACCGCGAG 2040
Db 2041 GCCAGGAAGCAGAAAGCAGATGCGGTGGCCACATGATGATGATGATGATGATGATG 2100
QY 2041 GCCAGGAAGCAGAAAGCAGATGCGGTGGCCACATGATGATGATGATGATGATGATG 2100
Db 2101 GGTGTAGCTTTGGCTTACGCTCTCTGCGGCGGCAAACTGGCAGCAACGCGTCTCAC 2160
QY 2101 GGTGTAGCTTTGGCTTACGCTCTCTGCGGCGGCAAACTGGCAGCAACGCGTCTCAC 2160
Db 2161 TTGATGCGCGCCACCGCAATATGATGATGATGATGATGATGATGATGATGATGATG 2220
QY 2161 TTGATGCGCGCCACCGCAATATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 2221 TGTGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280
QY 2221 TGTGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280
Db 2281 GCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
QY 2281 GCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db 2341 CCGGTACCAAAAGCAATGCGCAAAAGCAATGCGCAAAAGCAATGCGCAAAAGCAATG 2400
QY 2341 CCGGTACCAAAAGCAATGCGCAAAAGCAATGCGCAAAAGCAATGCGCAAAAGCAATG 2400
Db 2401 GGTTCGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
QY 2401 GGTTCGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
Db 2461 GCGAAGAGATCTCTGTTTAAAGGAGGTTCTACTGAGCAGCAGTGGCCCTGTTG 2520
QY 2461 GCGAAGAGATCTCTGTTTAAAGGAGGTTCTACTGAGCAGCAGTGGCCCTGTTG 2520
Db 2521 GCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
QY 2521 GCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
Db 2581 GCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
QY 2581 GCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2640
Db 2641 TAACTCCAAAATCCGGAAGGCTCTGTAATCCGGAAGGAGGAGGAGGAGGAGGAGGAGG 2700
QY 2641 TAACTCCAAAATCCGGAAGGCTCTGTAATCCGGAAGGAGGAGGAGGAGGAGGAGGAGG 2700
Db 2701 ACACACATACACAAAGAAAGAGTGGGTTCGCAAAATGTGAGAGAGAGCCAAAT 2760
QY 2701 ACACACATACACAAAGAAAGAGTGGGTTCGCAAAATGTGAGAGAGAGCCAAAT 2760
Db 2761 GTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820
QY 2761 GTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2820
Db 2821 ACTGTAACCTCCCTAAATAATTTGTATAGTAAATGAGCAAGCTGTGACCCAGCGTTTC 2880
QY 2821 ACTGTAACCTCCCTAAATAATTTGTATAGTAAATGAGCAAGCTGTGACCCAGCGTTTC 2880
Db 2881 GATCCCGAATTC 2892
QY 2881 GATCCCGAATTC 2892
RESULT 2

LOCUS 156055 2892 bp DNA PAT 12-AUG-1997
DEFINITION Sequence 5 from patent US 5648464.
ACCESSION 156055
NID 92476849
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2892)
AUTHORS Artavanis-Tsakonas,S., Fehon,R,Grant, Rebay,I. and
Blumuelier,C,Marie.
TITLE Human Notch and Delta binding domains in topolythmic proteins, and
JOURNAL methods based thereon
FEATURES
source location/Qualifiers
BASE COUNT 739 a 764 c 808 g 581 t
ORIGIN
Query Match 100.0%; Score 2892; DB 18; Length 2892;
Best Local Similarity 100.0%; Pred. No. 0.00e+00; Mismatches 0; Indels 0; Gaps 0;
Matches 2892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 GAATTCGAGGAGATTAATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 60
QY 1 GAATTCGAGGAGATTAATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTA 60
Db 61 CACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 120
QY 61 CACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 120
Db 121 GCACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 180
QY 121 GCACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 180
Db 181 TTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 181 TTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Db 241 ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 300
QY 241 ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 300
Db 301 ACAGGACAGTGCCTGGGACGCTGCAAGACGCGGTTTCCGCTGCTTAAGCACTACAG 360
QY 301 ACAGGACAGTGCCTGGGACGCTGCAAGACGCGGTTTCCGCTGCTTAAGCACTACAG 360
Db 361 GCGACATGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 420
QY 361 GCGACATGACACACACACACACACACACACACACACACACACACACACACACACACACACAC 420
Db 421 GAGAACTCGGTCAATCTGACGACGCGCCAGCGCTTCCAGAACAGGCTTCCAGATCC 480
QY 421 GAGAACTCGGTCAATCTGACGACGCGCCAGCGCTTCCAGAACAGGCTTCCAGATCC 480
Db 481 ATCCAGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 481 ATCCAGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 541 CATGATAGCAAAATAGCGGCAATGCGGCAATGCGGCAATGCGGCAATGCGGCAATGCGGCAAT 600
QY 541 CATGATAGCAAAATAGCGGCAATGCGGCAATGCGGCAATGCGGCAATGCGGCAATGCGGCAAT 600
Db 601 GTGACAGAGTACTGAGAGTCTCTCCGATGGAAGAGCAAGCAAGTCCGGAATCCGAGTAC 660
QY 601 GTGACAGAGTACTGAGAGTCTCTCCGATGGAAGAGCAAGCAAGTCCGGAATCCGAGTAC 660
Db 661 ACGTGCTGAGTACGATTTCCGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 661 ACGTGCTGAGTACGATTTCCGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 721 GCCAAGTTCTGCGGCGCGCGGACGACATTCATTTGACACTGACTTCTGCGAGAGCGGCG 780

RESULT 3
LOCUS 2889 bp RNA INV 30-NOV-1994
DEFINITION D.melanogaster DL mRNA for delta protein, involved in neurogenesis.
ACCESSION Y00222
NID 97835
KEYWORDS delta gene; neurogenesis.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryote; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 6 to 2884)
Kopczynski, C.C., Altom, A.K., Fechtel, K., Kooh, P.J. and
Muskavitch, M.A.
REFERENCE
AUTHORS Delta, a Drosophila neurogenic gene, is transcriptionally complex
and encodes a protein related to blood coagulation factors and
epidermal growth factor of vertebrates
Genes Dev. 2 (12B), 1723-1735 (1988)
JOURNAL 89196890
MEDLINE 2 (bases 1 to 2889)
AUTHORS Muskavitch, M.A.T.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1991) M.A.T. Muskavitch, Dept. of Biology,
Indiana University, Bloomington, Jordan Hall, A504, Indiana 47405,
USA
COMMENT See X06289 for overlapping sequence.
FEATURES
source
1. 2889
/location/Qualifiers
/organism="Drosophila melanogaster"
/strain="Oregon-R"
/db_xref="taxon:7227"
/haplotype="wildtype"
/dev_stage="4-7 hour embryo"
1. 2889
/gene="Dl"
/db_xref="FlyBase:FBgn0000463"
/evidence=experimental
1. 2889
/note="Delta"
/gene="Dl"
/allele=""
/db_xref="FlyBase:FBgn0000463"
142..2640
/gene="Dl"
/codon_start=1
/product="delta protein"
/db_xref="PID:g577774"
/translat="FBgn0000463"
/translation="MHWIKCLTAFICFTIVIVYHSSGSFELRLKYSNDHGRDNEGR
CCGESDGAATGKICSGCTRFKRLKATIDTSCGTGDTVPILGNSVYLDA
ORONKGTNPIDPFSESWPRTSLIVEMHNNNGNRNTNLLIORTLYOOLVY
SSEWTKNSOYTSLEDFRVTCDLNVHSGCAKPCRPBDSKSHGTCSEGTGIICT
TGOGDCHITPKCAKCEGHCDEKPNOCVOCLEKALCECVLEPNCIHGTCKRPT
CICNEGWSGLYCNODLNTCTNHRCKNGGCFNTEGLYCKAPGSGDCEYEITS
CDADVPCNGGTCIDEPHTKTKCKCANGWSGKMEKRYLTCSDPCHQDCEVETIS
PGLSGKGYCCCEPIGSGPNCIDLONCSNPCIISGSCQPSGKICPAGESGTRC
ETNIDICIGHGCGNGTCIDMNOYRCOCVCPGFGTHSSKVDCLIRPANGGTCIN
LNDVYCTCRASTGKDCSYDIDECSSGPRHNGTCNNRNSPECVANCPRKQDCE
ESTDSTVFDAHQGTATTOARADGLTNQVVLVAFVAMLEAVIAVCVFKRKKK
RAQEKDAEARKQNEONAIVATMHNHNSGVVALASALSGKTSNSLTFDGGNPII
KNTWDSVNNICAAALAAAAAADCLMAGTGVAVADNNNNANSDFCVAPLORAS
OKOINTDPTLMHRSAGSAGKAGSGGSGAABEGRISVLGEGSYCSQWPSLAAGV
AGACSSOLMAAASAAAGTGTAAQOQRSSVYCGTTPH"
142..213
/gene="Dl"
/db_xref="FlyBase:FBgn0000463"
214..2640
/gene="Dl"
/db_xref="FlyBase:FBgn0000463"
/product="delta protein"
820..1839
repeat_region

misc_feature
1999..2016
/note="cysteine-rich, factor IX/EGF-like repeats"
/gene="Dl"
/note="stop-transfer sequence"
/db_xref="FlyBase:FBgn0000463"
BASE COUNT 739 a 764 c 805 g 581 t
ORIGIN
Query Match 99.3%; Score 2871; DB 15; Length 2889;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2889; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Db 1 GAATTCGAGGAGATTTATCAAAACATAAACAATTAAGTGTAGTGGCCGACA 60
Qy 1 GAATTCGAGGAGATTTATCAAAACATAAACAATTAAGTGTAGTGGCCGACA 60
Db 61 CACACACACACACAGCCCGGTGATTTATCACTAAACGCACTCAATCAAAAAATCA 120
Qy 61 CACACACACACACAGCCCGGTGATTTATCACTAAACGCACTCAATCAAAAAATCA 120
Db 121 GCACAAAACATCATTAATCATGATTTGATTAATTTTATTAACAGCATTCATTGC 180
Qy 121 GCACAAAACATCATTAATCATGATTTGATTAATTTTATTAACAGCATTCATTGC 180
Db 181 TTCACAGTCATGTCAGGTCACAGTCCCGGACCTTGAGTGGCGCTGAAGTCTTC 240
Qy 181 TTCACAGTCATGTCAGGTCACAGTCCCGGACCTTGAGTGGCGCTGAAGTCTTC 240
Db 241 AGCAACGATCACAGGCGGAGACAGAGGCTGCTGTCACAGGAGATCGAGACG 300
Qy 241 AGCAACGATCACAGGCGGAGACAGAGGCTGCTGTCACAGGAGATCGAGACG 300
Db 301 ACGGCAAGTGCCTGGGACGCTGCAAGACGGGTTTCCGCTGCTTAAGCATCTACAG 360
Qy 301 ACGGCAAGTGCCTGGGACGCTGCAAGACGGGTTTCCGCTGCTTAAGCATCTACAG 360
Db 361 GCCACCATGACACACCTCCAGTGCCTACAGGAGGAGATCAGCCCATCTCTGGC 420
Qy 361 GCCACCATGACACACCTCCAGTGCCTACAGGAGGAGATCAGCCCATCTCTGGC 420
Db 421 GAGAACTGGTCAATCTGACGAGCCGACGCTCCGAAACAAAGGCTTACAGATCC 480
Qy 421 GAGAACTGGTCAATCTGACGAGCCGACGCTCCGAAACAAAGGCTTACAGATCC 480
Db 481 ATCCAGTCCCTCTCTCTCATGAGCGGGGTACCTTCTGCTGANTGCTGAGGCTGG 540
Qy 481 ATCCAGTCCCTCTCTCTCATGAGCGGGGTACCTTCTGCTGANTGCTGAGGCTGG 540
Db 541 CATGATAGCAATAGCGGCAATGCGGAAACCAAGCTCTCATCCAGGACTCTTG 600
Qy 541 CATGATAGCAATAGCGGCAATGCGGAAACCAAGCTCTCATCCAGGACTCTTG 600
Db 601 GTGCGACGAGTACGAGAGTGTCTCCGAATGGAAGCAAGTGGGAATCGGAGTAC 660
Qy 601 GTGCGACGAGTACGAGAGTGTCTCCGAATGGAAGCAAGTGGGAATCGGAGTAC 660
Db 661 ACCTGCTGAGAGTACGATTTCCGCTGTCACCTGCGATCTCAACTAGTACGAGTCCG 720
Qy 661 ACCTGCTGAGAGTACGATTTCCGCTGTCACCTGCGATCTCAACTAGTACGAGTCCG 720
Db 721 GCGAAGTCTGCGGCGCGCGGAGATTCATTTGGACACTGACTTGTGCGAAGCGGGC 780
Qy 721 GCGAAGTCTGCGGCGCGCGGAGATTCATTTGGACACTGACTTGTGCGAAGCGGGC 780
Db 781 GAAATATCTGTTGACCGGATGGAGGGGCTTACTGTCATACCAATGCGCCAAA 840
Qy 781 GAAATATCTGTTGACCGGATGGAGGGGCTTACTGTCATACCAATGCGCCAAA 840
Db 841 GGCTGTGAACATGACATTTGGACAAACCAATCAATGCTGTTGGCAACTGGGCTGAG 900
Qy 841 GGCTGTGAACATGACATTTGGACAAACCAATCAATGCTGTTGGCAACTGGGCTGAG 900
Db 901 GGAGCTTGTCAACGAGTGGCTTGTGACCCGAATGATCATGACCTGCAACAA 960

|||||
QY 901 GGAGCCTGTGCAAGAGTGGCTTCTGNACCGAACTGCATCCATGGCACTGCAACAA 960
Db 961 CCCGGACTTGCATCTGCAAGAGAGGTTGGGAGGCTTGTACTGCAACACGAGATCTGAAC 1020
QY 961 CCCGGACTTGCATCTGCAACAGAGGTTGGGAGGCTTGTACTGCAACACGAGATCTGAAC 1020
Db 1021 TACTGCACCAACACAGACCTTGCAGAGATGGCGGAACTGTTCAACACCGCGGAGGA 1080
QY 1021 TACTGCACCAACACAGACCTTGCAGAGATGGCGGAACTGTTCAACACCGCGGAGGA 1080
Db 1081 TTGTACACATGCAATGCGCTCCAGATACAGTGTGATGTTGGAAATGAGATCTAC 1140
QY 1081 TTGTACACATGCAATGCGCTCCAGATACAGTGTGATGTTGGAAATGAGATCTAC 1140
Db 1141 TCCTGCGATGCGGATGTCAATCCCTGCCAGAAATGGTGTACCTGCATCGATGAGCGCAC 1200
QY 1141 TCCTGCGATGCGGATGTCAATCCCTGCCAGAAATGGTGTACCTGCATCGATGAGCGCAC 1200
Db 1201 AAAAAACCGGCTACAAGTGTCAATTTGCCCAACGCTGGAGCGGAAAGATGTGCGAGGAG 1260
QY 1201 AAAAAACCGGCTACAAGTGTCAATTTGCCCAACGCTGGAGCGGAAAGATGTGCGAGGAG 1260
Db 1261 AAAGTGTCTACGCTGTTCGGACAAACCTGTCTATCAGGGAATCTGCCGCAACGTTGCTCCT 1320
QY 1261 AAAGTGTCTACGCTGTTCGGACAAACCTGTCTATCAGGGAATCTGCCGCAACGTTGCTCCT 1320
Db 1321 GGCTTGGGAAGCAAGGTCAGGGCTACCAAGTGCAGATGTCGAATGTCCTATGGCTACAGCGGACCC 1380
QY 1321 GGCTTGGGAAGCAAGGTCAGGGCTACCAAGTGCAGATGTCGAATGTCCTATGGCTACAGCGGACCC 1380
Db 1381 AACTGCGATCTCCAGCTGGACAACTGCAGTCCGAATCCATGATTAACCGGTGGAAGCTGT 1440
QY 1381 AACTGCGATCTCCAGCTGGACAACTGCAGTCCGAATCCATGATTAACCGGTGGAAGCTGT 1440
Db 1441 CAGCGAGCGGAAGTGTATTTGCCAGCGGATTTTCGGGAACGAGATGCGGAGACCAAC 1500
QY 1441 CAGCGAGCGGAAGTGTATTTGCCAGCGGATTTTCGGGAACGAGATGCGGAGACCAAC 1500
Db 1501 ATTGACGATGTCTTGGCCACAGCTGCGAGAACGAGGACCTGCTATAGATATGTTCAAC 1560
QY 1501 ATTGACGATGTCTTGGCCACAGCTGCGAGAACGAGGACCTGCTATAGATATGTTCAAC 1560
Db 1561 CAATATCGTGCATATGCTTCCCGTTCCATGGCACCCACTGTAGTAGCAAGTTGAC 1620
QY 1561 CAATATCGTGCATATGCTTCCCGTTCCATGGCACCCACTGTAGTAGCAAGTTGAC 1620
Db 1621 TTGTGCTCATCAGACCGTGTGCCAATGAGGAACCTGCTTGAATCTCAACACGATTAC 1680
QY 1621 TTGTGCTCATCAGACCGTGTGCCAATGAGGAACCTGCTTGAATCTCAACACGATTAC 1680
Db 1681 CAGTGCACCTGTCTGCGGGATTTACTTGGCAAGGATGCTGTGGACATCGATGAGTGC 1740
QY 1681 CAGTGCACCTGTCTGCGGGATTTACTTGGCAAGGATGCTGTGGACATCGATGAGTGC 1740
Db 1741 AGCAGTGACCTGTCTATTAACCGGCGCACTTGCATGAACCGCGTCAATTCGTTGCGAATGC 1800
QY 1741 AGCAGTGACCTGTCTATTAACCGGCGCACTTGCATGAACCGCGTCAATTCGTTGCGAATGC 1800
Db 1801 GTGTGTGCCAATGTTTTCAGGGGCAAGCAGTGCATGAGGAGTCCCTACGATTCGTTGACC 1860
QY 1801 GTGTGTGCCAATGTTTTCAGGGGCAAGCAGTGCATGAGGAGTCCCTACGATTCGTTGACC 1860
Db 1861 TTCGATGCCCAATATATGAGCGCACCAACAGCGAGAGCGGATGTTGACCAATGCC 1920
QY 1861 TTCGATGCCCAATATATGAGCGCACCAACAGCGAGAGCGGATGTTGACCAATGCC 1920
Db 1921 CAGGTAGTCTTAATGCTGTTTCTCCGTTGCGATGCCCTTTGTTGCGGTTATTCGCGGG 1980
QY 1921 CAGGTAGTCTTAATGCTGTTTCTCCGTTGCGATGCCCTTTGTTGCGGTTATTCGCGGG 1980
Db 1981 TSCGTGGTCTTCTGCATGAAGCGCAAGCCTAAGCGTGTCTCAGAAAGGACGACGCGGAG 2040
|||||

QY 1981 TCSCGTGGTCTTCTGCATGAAGCGCAAGCGTAAGCGTGTCTCAGAAAAAGGACGACGCGGAG 2040
Db 2041 GCCAGAGCAGAAACAGACAGAAATCGGTGGCCACAATGCAATCAACAATGGCGAGTGGGTG 2100
QY 2041 GCCAGAGCAGAAACAGACAGAAATCGGTGGCCACAATGCAATCAACAATGGCGAGTGGGTG 2100
Db 2101 GGTGTAGCTTTTGGCTTTCAGCCTCTCTGGCGGCAAAACCTGGCAGCAACAGCGGTCTCAC 2160
QY 2101 GGTGTAGCTTTTGGCTTTCAGCCTCTCTGGCGGCAAAACCTGGCAGCAACAGCGGTCTCAC 2160
Db 2161 TTCGATGGCGCAACCCGGAATATCATCAAAAAACCTGGGACAAGTTCGGTCAACAACATT 2220
QY 2161 TTCGATGGCGCAACCCGGAATATCATCAAAAAACCTGGGACAAGTTCGGTCAACAACATT 2220
Db 2221 TGTGCTCAGCAGCAGCAGCGGCGGCGGAGCAGCAGCGGCGAGAGTGTCTCATGTAC 2280
QY 2221 TGTGCTCAGCAGCAGCAGCGGCGGCGGAGCAGCAGCGGCGAGAGTGTCTCATGTAC 2280
Db 2281 GCGGATATGTGCTTTCGGTGGGATACAAACATGCCAACTCAGACTTTTGTGTGGCT 2340
QY 2281 GCGGATATGTGCTTTCGGTGGGATACAAACATGCCAACTCAGACTTTTGTGTGGCT 2340
Db 2341 CCGCTACAAAGAGCCAAAGTCGCAAAAGCAACTCAACACCGATCCACGCTCATGCACGC 2400
QY 2341 CCGCTACAAAGAGCCAAAGTCGCAAAAGCAACTCAACACCGATCCACGCTCATGCACGC 2400
Db 2401 GGTTCGCGGAGGAGCTCAGCAGGAGGCTGTGGCGGAGGACCGGCGGCGGAG 2460
QY 2401 GGTTCGCGGAGGAGCTCAGCAGGAGGCTGTGGCGGAGGACCGGCGGCGGAG 2460
Db 2461 GCACAGAGATCTCTGTTTTCGGTGGGAGGCTTCTACTGTAGCCAGGTTGGCCCTCGTTG 2520
QY 2461 GCACAGAGATCTCTGTTTTCGGTGGGAGGCTTCTACTGTAGCCAGGTTGGCCCTCGTTG 2520
Db 2521 GCGGCGGCGGAGTGGCGGAGGCTTCTCATCCAGCTAATGGCTGACGCTTCGCGCAGCG 2580
QY 2521 GCGGCGGCGGAGTGGCGGAGGCTTCTCATCCAGCTAATGGCTGACGCTTCGCGCAGCG 2580
Db 2581 GGCA-CGGG-CGGG-ACGGCGCAACAGCAGATCGTGGTCTGGGCACTCCGCAATATG 2637
QY 2581 GGCAGCGGAGCGGCGGCGGCAACAGCAGATCGTGGTCTGGGCACTCCGCAATATG 2640
Db 2638 TTAATCCCAAAATCCGGAAGGCTCTCTGTAATATCCGGAATAATCCGATGAGGAGCTG 2697
QY 2641 TTAATCCCAAAATCCGGAAGGCTCTCTGTAATATCCGGAATAATCCGATGAGGAGCTG 2700
Db 2698 ACAGCACAATACAAAGAAAGAACTGGGTTGGGTTCAAAATGTGAGAGAGAGCCCAAAAT 2757
QY 2701 ACAGCACAATACAAAGAAAGAACTGGGTTGGGTTCAAAATGTGAGAGAGAGCCCAAAAT 2760
Db 2758 GTTGTGTTGATTGAAGCAGTGTAGTCTCAGCAAAAAATGAAAAATCTGTAAACAGCATA 2817
QY 2761 GTTGTGTTGATTGAAGCAGTGTAGTCTCAGCAAAAAATGAAAAATCTGTAAACAGCATA 2820
Db 2818 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAACTGTGACCCAGCGGTTTC 2877
QY 2821 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAACTGTGACCCAGCGGTTTC 2880
Db 2878 GATCCCGAATTC 2889
QY 2881 GATCCCGAATTC 2892

RESULT 4
LOCUS Drosophila mRNA for Delta (D1) gene.
DEFINITION X06289
ACCESSION 97852
NID
KEYWORDS Drosophila melanogaster
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE 1 (bases 1 to 4724)
 AUTHORS Campos-Ortega, J.A.
 TITLE Direct Submission
 JOURNAL Submitted (24-NOV-1987) Campos-Ortega J.A., Institut fuer
 Entwicklungsphysiologie, Gyrohofstr. 17, D-5000 Cologne, FRG
 REFERENCE 2 (bases 1 to 4724)
 AUTHORS Vaessen, H., Bremer, K.A., Knust, E. and Campos-Ortega, J.A.
 TITLE The neurogenic Gene Delta of Drosophila melanogaster is expressed
 in neurogenic territories and encodes a putative transmembrane
 protein with EGF-like repeats
 JOURNAL EMBO J. 6, 3431-3440 (1987)
 COMMENT include x05140 (referring to the authors); some differences found.
 FEATURES
 source location/Qualifiers
 1. 4724
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /dev_stage="embryonic"
 /clone_1lb="CDNA in EMBL4"
 /clone="C3.2"
 /map="chromosome 3 band 92A2"
 127. 2769
 /note="Delta"
 /gene="Dl"
 /db_xref="FlyBase:FBgn0000463"
 127. 180
 /gene="Dl"
 /db_xref="FlyBase:FBgn0000463"
 127. 2769
 /gene="Dl"
 /note="delta Dl"
 /codon_start=-1
 /db_xref="PID:97853"
 /db_xref="SWISS-PROT:P10041"
 /translation="MHVINCILAFICFYIYVHSGSEELKRYFSDNHDNEER
 CCSGESDGAATGKCLGSKCTFRCLKHYQATIDTTSCTGIDVITPLIGNSVNLDA
 ORONKGFNPIDPFPSFWPTESLIVEMHDTNNSGNARKLQLRLVLOVLE
 SEMKTKNESQYTSLEDFRYVCDIYSGGACFCRPRDSDGSHGSEIETL
 TGMQGDYCHIPKDCGCEHCHDKPNOVCVOLGKMGALNECVLEPNCIHGTENKPT
 CINCNEGGLCYCNDKLCYCNHPRCKCFMDEGLTYCKAPGSGDCEMEIYS
 CDADVPCONGTICIDEPRHTKTYKCHRNKMSGKMEKVEVLYTCSBPQOICRNYR
 PEGSGSGQYQCEPCIGYSGPNCIDLDNCSNPNCINGSCQPSGKCTCPSGSTGC
 ETNIDICLHOCENGCTCIDMNOYRQOCVPHGTHCSKVDLCILRPGANGSTCN
 LNDYDCTCRAGFTGKDCSYDIDESGSPCHNGCTCMRNVSFEVCANRFRKQCE
 ESTDSTFDHOGVATTTOARADGLANAOVYLIVFSVAMPVLAVIACVYFCRRKX
 RADERNABARKRONONAVATTHHNSAVGALASAMGKTSNSGLTDCGNPTI
 KNTWDSVNNICASAAAAAADADECLTGTIVASADNNNNNSDFCAPLQRAKS
 OKQNLNTPILMERGSPAGTSAGKASGGGPGAPGKRLSYLDESGSYCSQMPSLAAAV
 AGDLFIOLMAASVAGTGTACQOORSYVCTPHYVPIKIRKSGMEIRNPHGADSTY
 TKRLGMVONVRETPKMLLLEAV"
 181. 2766
 /gene="Dl"
 /db_xref="FlyBase:FBgn0000463"
 75. 1836
 /gene="Dl"
 /note="region of 10 EGF-like repeats"
 /db_xref="FlyBase:FBgn0000463"
 4705. 4711
 /note="polyA_signal"
 4724
 BASE COUNT 1465 a 1073 c 1045 g 1141 t
 ORIGIN
 Query Match 95.08; Score 2746; DB 15; Length 4724;
 Best Local Similarity 98.58; Pred. No. 0.00e+00;
 Matches 2829; Conservative 0; Mismatches 35; Indels 8; Gaps 8;

Db 1 TTGAAAACATTAACATTAACATTTGAGTAGTGGCCGACACACACACACACAG 60
 Oy 17 TTGAAAACATTAACATTAACATTTGAGTAG-TTGCCGACACACACACACACAG 75
 Db 61 CCCGTGATATTACTAAAGGACACATCAACCAAAAATCAGCAACAAAATCA 120
 Oy 76 CCCGTGATATTACTAAAGGACACACATCAACCAAAAATCAGCAACAAAATCA 135
 Db 121 ATAAACATGATGATTAATTAATTAATTAACAGCATTCATTTGCTTACAGTATCGTG 180
 Oy 136 ATAAACATGATGATTAATTAATTAATTAACAGCATTCATTTGCTTACAGTATCGTG 195
 Db 181 CAGGTTCACATTTCCGGGACGCTTGTAGTGGCCCTGAAGTACTTTCAGCAAGATCATGGG 240
 Oy 196 CAGGTTCACATTTCCGGGACGCTTGTAGTGGCCCTGAAGTACTTTCAGCAAGATCATGGG 255
 Db 241 CGGGCAACGAGGGGCTGTGTGACGAGGGGAGTGGACGGGAGGAGGAGGAGGAGGCTG 300
 Oy 256 CGGGCAACGAGGGGCTGTGTGACGAGGGGAGTGGACGGGAGGAGGAGGAGGAGGCTG 315
 Db 301 GGCAGCTGCAAGAGCGGCTTTCGGCTCTGCTTAAGCACTACCAAGGCCACCATGCACAC 360
 Oy 316 GGCAGCTGCAAGAGCGGCTTTCGGCTCTGCTTAAGCACTACCAAGGCCACCATGCACAC 375
 Db 361 ACCTCCAGTGCACCTACGCGGAGCGTATACAGCCCATTTCTCGGAGACTCGGTCAAT 420
 Oy 376 ACCTCCAGTGCACCTACGCGGAGCGTATACAGCCCATTTCTCGGAGACTCGGTCAAT 435
 Db 421 CTGACCGAGCGCCAGCGCTTCCAGAACAGGGCTTCACAGATCCCATTCAGTCCCTTC 480
 Oy 436 CTGACCGAGCGCCAGCGCTTCCAGAACAGGGCTTCACAGATCCCATTCAGTCCCTTC 495
 Db 481 TCGTTTCATGCGCGGAGTACCTTCTCGCTGATCGTGAAGGCTGCGATGATACGAACAC 540
 Oy 496 TCGTTTCATGCGCGGAGTACCTTCTCGCTGATCGTGAAGGCTGCGATGATACGAACAT 555
 Db 541 AGCGGCAATGCGCGGAGTACCTTCTCGCTGATCGTGAAGGCTGCGATGATACGAACAT 600
 Oy 556 AGCGGCAATGCGCGGAGTACCTTCTCGCTGATCGTGAAGGCTGCGATGATACGAACAT 615
 Db 601 GAGGTGCTCTCGATGGAAGAACAGCAAGTGGAAATGCGAGTACAGTCCCTGGAGTAC 660
 Oy 616 GAGGTGCTCTCGATGGAAGAACAGCAAGTGGAAATGCGAGTACAGTCCCTGGAGTAC 675
 Db 661 GATTTCGCTGACCTGCGATCTCACTACGATACGATCCGCTGTCGCAATTTCTGCGCG 720
 Oy 676 GATTTCGCTGACCTGCGATCTCACTACGATACGATCCGCTGTCGCAATTTCTGCGCG 735
 Db 721 CCCGCGAGATGATTTGACACTGCACTTCTCGGAGAGCGGCGAAATATCTGTTTG 780
 Oy 736 CCCGCGAGATGATTTGACACTGCACTTCTCGGAGAGCGGCGAAATATCTGTTTG 795
 Db 781 ACCGATGCGAGGCGATTTGATGATACATCCCAATGCGGCAAAAGCTGTGAACATGGA 840
 Oy 796 ACCGATGCGAGGCGATTTGATGATACATCCCAATGCGGCAAAAGCTGTGAACATGGA 855
 Db 841 CATTGCGACAGCCCAATCAATGCTTGGCAACTGGGCTGGAAGAGGCTTGTGCAAC 900
 Oy 856 CATTGCGACAGCCCAATCAATGCTTGGCAACTGGGCTGGAAGAGGCTTGTGCAAC 915
 Db 901 GAGTGCCTTGTGAACCGCACTGATCATGCACTGCAACCAACCTGTGACTTGCATC 960
 Oy 916 GAGTGCCTTGTGAACCGCACTGATCATGCACTGCAACCAACCTGTGACTTGCATC 975
 Db 961 TGCACAGAGGATGGGAGGCTTGTACTGCAACAGAGTTTGAATCTGACCAACACAC 1020
 Oy 976 TGCACAGAGGATGGGAGGCTTGTACTGCAACAGAGTTTGAATCTGACCAACACAC 1035
 Db 1021 AGACCTGCAAGATGGGAGGCTTGTACTGCAACAGAGTTTGAATCTGACCAACACAC 1080
 Oy 1036 AGACCTGCAAGATGGGAGGCTTGTACTGCAACAGAGTTTGAATCTGACCAACACAC 1095
 Db 1081 TGGCTCCAGAGATACAGTGTGATGATTTGCAAAATGAGATCTACTCTGCGATGCCGAT 1140
 Oy 1096 TGGCTCCAGAGATACAGTGTGATGATTTGCAAAATGAGATCTACTCTGCGATGCCGAT 1155
 Db 1141 GTCAATCCCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200

QY 1156 GTCAATCCTCGCAGAAATGGTGGTACCTCATCGATGAGCCGCACACAAAACCGGCTAC 1215
Db 1201 AAGTGTCAATTCGCGAACCGCTGGAGCGGAAAGATGTGCGAGGAGAAAGTGTCTCAGTGT 1260
QY 1216 AAGTGTCAATTCGCGAACCGCTGGAGCGGAAAGATGTGCGAGGAGAAAGTGTCTCAGTGT 1275
Db 1261 TCGGACAAACCTGTCTCATCAGGGAATCTGCCGCAACGTTGCTCTGTTGGGAACAAG 1320
QY 1276 TCGGACAAACCTGTCTCATCAGGGAATCTGCCGCAACGTTGCTCTGTTGGGAACAAG 1335
Db 1321 GGTCAAGGCTACAGTGCAGTCCCAATTGCTTGGCTACAGCGGACCCCAACTGCGATCTCCAG 1380
QY 1336 GGTCAAGGCTACAGTGCAGTCCCAATTGCTTGGCTACAGCGGACCCCAACTGCGATCTCCAG 1395
Db 1381 GTGGAACCTGCAAGTCCCAATCCCTGCATTAAGCGTGGAGAGTGTGACGCGGCGGAAAG 1440
QY 1396 GTGGAACCTGCAAGTCCCAATCCCAATTAAGCGTGGAGAGTGTGACGCGGCGGAAAG 1455
Db 1441 TGTATTTGGCCAGCGGATTTTCGGGAAGAGGTGCGAGACCAACATTTGACGATTTGCTT 1500
QY 1456 TGTATTTGGCCAGCGGATTTTCGGGAAGAGATGCGAGACCAACATTTGACGATTTGCTT 1515
Db 1501 GCGCACCACTGCGAGAACGAGGACCTGCATAGATATGTCACCAATATGCTGCGCAA 1560
QY 1516 GCGCACCACTGCGAGAACGAGGACCTGCATAGATATGTCACCAATATGCTGCGCAA 1575
Db 1561 TCGGTTCCCGGTTTCCATGGCACCCACTGTAGTAGCAAAAGTTGACTTGTGCTTCATCAGA 1620
QY 1576 TCGGTTCCCGGTTTCCATGGCACCCACTGTAGTAGCAAAAGTTGACTTGTGCTTCATCAGA 1635
Db 1621 CCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACACGATTTACGAGTGCACTGTCTGT 1680
QY 1636 CCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACACGATTTACGAGTGCACTGTCTGT 1695
Db 1681 GCGGATTTACTGGCAAGGATTCCTCGTGGACATCGATGAGTGACAGTGACACCTGT 1740
QY 1696 GCGGATTTACTGGCAAGGATTCCTGTGGACATCGATGAGTGACAGTGAGTGACACCTGT 1755
Db 1741 CATAAGCGCGCACTTGCATGAACCGCGTCAATTCGTTGCAATGCGTGTGCGCAATGGT 1800
QY 1756 CATAAGCGCGCACTTGCATGAACCGCGTCAATTCGTTGCAATGCGTGTGCGCAATGGT 1815
Db 1801 TTCAGGGCAAGCAGTGCATGAGGAGTCTACGATTCGGTGACCTTCGATGCCACCAA 1860
QY 1816 TTCAGGGCAAGCAGTGCATGAGGAGTCTACGATTCGGTGACCTTCGATGCCACCAA 1875
Db 1861 TATGAGCGCACACAAAGCAGAGCGGATGTTTAGCCCAATCCCGAGGAGTCCCTAAT 1920
QY 1876 TATGAGCGCACACAAAGCAGAGCGGATGTTTAGCCCAATCCCGAGGAGTCCCTAAT 1935
Db 1921 GCTGTTTCTCGGTTGCGATGCTTTGGTGGCGGTTATTGCGCGTGCCTGCTTCTGC 1980
QY 1936 GCTGTTTCTCGGTTGCGATGCTTTGGTGGCGGTTATTGCGCGTGCCTGCTTCTGC 1995
Db 1981 ATGAGCGCAAGGTAAAGGTGCTAGGAAAGAGCAACCGGAGGCGCAGGAGAGCAAC 2040
QY 1996 ATGAGCGCAAGGTAAAGGTGCTAGGAAAGAGCAACCGGAGGCGCAGGAGAGCAAC 2055
Db 2041 GAACAGAAATGCAGTGCATCAATGCGAGTGCCTGCGTGGGTGAGCTTTGGCT 2100
QY 2056 GAACAGAAATGCAGTGCATCAATGCGAGTGCCTGCGTGGGTGAGCTTTGGCT 2115
Db 2101 TCAGCCTCTATGGCGCGCAAACTGCGAGCAACAGCGGTTCTACCTTCGATGCGGCAAC 2160
QY 2116 TCAGCCTCTATGGCGCGCAAACTGCGAGCAACAGCGGTTCTACCTTCGATGCGGCAAC 2175
Db 2161 CCGAATATCATCAAAACACCTGGGACAAAGTCCGTCACCAACAAATTTGTGCTCAGCAGCA 2220
QY 2176 CCGAATATCATCAAAACACCTGGGACAAAGTCCGTCACCAACAAATTTGTGCTCAGCAGCA 2235
Db 2221 GCAGCGCGCGCGCGCAGCAGCGGCGGAGTGTCTCATGTACGCGGATATGTGCC 2280
QY 2236 GCAGCGCGCGCGCGCAGCAGCGGCGGAGTGTCTCATGTACGCGGATATGTGCC 2295

Db 2281 TCGGTGGCGGATTAACAACAATGCCAACTCAGACTTTTGTGTGCTCCGCTACAAAAGGCC 2340
QY 2296 TCGGTGGCGGATTAACAACAATGCCAACTCAGACTTTTGTGTGCTCCGCTACAAAAGGCC 2355
Db 2341 AAGTCCAAAAGCAACTCAACACCGATCCACGCTCATGCTCGCGTTCGCGCGGAGGC 2400
QY 2356 AAGTCCAAAAGCAACTCAACACCGATCCACGCTCATGCTCGCGTTCGCGCGGAGGC 2415
Db 2401 ACGTCCGCCAAAGGAGCGTCTGGCGGAGGACCGGAGCGCGGAGGCAAGAGATCTCC 2460
QY 2416 AGCTACGCCAAAGGAGCGTCTGGCGGAGGACCGGAGCGCGGAGGCAAGAGATCTCT 2475
Db 2461 GTTTTAGCGAGGTTCTCTACTGTAGCCAGCTTGCCCTCTGTTGGCGGCGGAGTG 2520
QY 2476 GTTTTAGCGAGGTTCTCTACTGTAGCCAGCTTGCCCTCTGTTGGCGGCGGAGTG 2535
Db 2521 GCCGGAGACTGTTCCATCCAA-CTAATGGCTGCAGCTTCGGTAGCGGCA-CGGA-CGGG 2577
QY 2536 GCCGGAG-CTGTTCATCCAGCTAATGGCTGCAGCTTCGGCAGCGGCGAGCGGCGG 2594
Db 2578 -ACGGCGCAACAGCAGCATCCGTGCTCGCGCACTCCCGCATTTATGTAACCTCAAAAAT 2636
QY 2595 GACGGCGCAACAGCAGGATCCGTGCTCGCGCACTCCCGCAT-ATGTAACCTCAAAAAT 2653
Db 2637 CCGGAAGGCTCTCGGAATCCGGAATCCGGAATCCGATGGAGAGCTGACAGCACAATAC 2696
QY 2654 CCGGAAGGCTCTCGTAAATCCGGAATCCGATGGAGAGCTGACAGCACAATAC 2713
Db 2697 AAGAAAGACTGGTGGTTCAAAATGTGAGAGAGCGCCCAAAATGTTGTTGTAT 2756
QY 2714 AAGAAAGACTGGTGGTTCAAAATGTGAGAGAGCGCC-AAAATGTTGTTGTAT 2772
Db 2757 TGAAGCAGTTTACTGCTACGAAAATGAAAATCTGTACAGGCACTAATCTGAAATC 2816
QY 2773 TGAAGCAGTTTACTGCTACGAAAATGAAAATCTGTACAGGCACTAATCTGAAATC 2832
Db 2817 CTTAAAAATTTGTATAGTAATTTAGCAAGCTGTGACCCAGCGGTTTCGATC 2868
QY 2833 CTTAAAAATTTGTATAGTAATTTAGCAAGCTGTGACCCAGCGGTTTCGATC 2884

5

RESULT 5
LOCUS DMDLEGFH 600 bp RNA INV 04-AUG-1995
DEFINITION Drosophila melanogaster mRNA fragment of D1 locus with EGF-like repeats (EGF= epidermal growth factor).
ACCESSION X05140
NID 97851
KEYWORDS D1 gene; EGF-like sequence; epidermal growth factor; neurogenesis; unidentified reading frame.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda; Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 600)
AUTHORS Knust,E., Dietrich,U., Tepass,U., Bremer,K.A., Weigel,D., Vassin,H. and Campos-Ortega,J.A.
TITLE EGF homologous sequences encoded in the genome of Drosophila melanogaster, and their relation to neurogenic genes
JOURNAL EMBO J. 6 (3), 761-766 (1987)
MEDLINE 87218537
FEATURES
Location/Qualifiers
source 1. .600
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
misc_feature 1. .447
/gene="D1"
/note="4x repeated unit"
/db_xref="FlyBase:FBgn0000463"
gene 1. .600
/note="Delta"
/gene="D1"

| | | | |
|-----------|---------------------------------------------------------------|-----------------------------------------------------------------|------------------------------|
| QY | 496 | TCGTTCTCATGCGGGTACCTTCTCGCTGATGCTCGAGCG-CTGG-CAT-GATAGAAC | 552 |
| Db | 418 | GACGACCTCACACAGAAACCCGAGCGCTCATCAGCGCGCTGCGCCACCCAGAGGCAC | 477 |
| QY | 553 | AATAGCGCAATGCGGGAACCAAGCTCCTCATCCAGGACTCTTGGTGCAGAGGTA | 612 |
| Db | 478 | CTGGGGTGGGAGAGGTGGTCCAGGACCTGCACAGAGCGCGCCGACCCACCTCAAG | 537 |
| QY | 613 | CTGAGAGTCTCTCCGAATGGAAGACCAAGCTCGGAATCGCAGTACACGTCGCTGGAG | 672 |
| Db | 538 | TACTCTATPCTGTTGTGTGATGAGCACTACTACGGGAAGCTGCTCTGCTCTGTC | 597 |
| QY | 673 | TACGATTCCGTGTACCTTGGCATCTCACTACTACGATCCGCGCTGTGCCAAGTCTCTG | 732 |
| Db | 598 | CGGCCCCGTGACGACCGCTTCGGTCACTTCCAGCTGTGGAGAGCGTGGCGAGAAGTCTGC | 657 |
| QY | 733 | CGGCCCCGCGAGATTCAATTGGACACTCGACTTGTCTCGAGACGGCGGAATTAATCTGT | 792 |
| Db | 658 | AACCCAGGTGGAAGGGCCAGTACTGCTGCTGAGCCGATTTGCTTGCTGCGGTGTGAGGAG | 717 |
| QY | 793 | TTGACCGGATGGCAGGGCGATTACTGTCACTACCAACCAATGCCCAAGGCTGTGA--A- | 849 |
| Db | 718 | CAGCAGCGCTTCTGCACAAACCTGGGAATGCAAGTGCAGAGTGGTGGCAGGGCGG | 777 |
| QY | 850 | CATG--GACAT-TGCAGAAACCAATCAATGCTTGGCACTGGCTGGGAAGGAGCC | 906 |
| Db | 778 | TACTGTGACGATGTCATCCGATACCCAGGCTGCTGCACGGTACCTGTCAAGCAGGACTGTC | 837 |
| QY | 907 | TTGTGCAAGGATGCTTCTGGAACCGAACTGCATCATCCATGGCAGCTGCAACAACCTGG | 966 |
| Db | 838 | CAGTGAACCTGCGAAGAGCTGGGCGGCTTTCTGCAACAGGACCTGAACTACTGTC | 897 |
| QY | 967 | ACTTGCACTGCAACAGAGGTTGGGAGGCTTGTACTGCAACAGGATCTGAACTACTGC | 1026 |
| Db | 898 | ACTCACCACAGCATGCAAGATGTGCAATGTCACATGACCAACACCGTCAAGGGAGCTAC | 957 |
| QY | 1027 | ACCAACACAGACCTTGCAGAAATGCGGNACTGTTCAACACCGCGAGGATGTATC | 1086 |
| Db | 958 | ACTGTTCTTCCGACCTGGGTACACAGGCTCCAGCTGCGAGATTGAAATCAACGAATGT | 1017 |
| QY | 1087 | ACATCAAAATGCTCCAGGATACATGGTGATGATTGCGAAAATGAGATCTACTCTGTC | 1146 |
| Db | 1018 | GATGCC 1023 | |
| QY | 1147 | GATGCC 1152 | |
| RESULT | 7 | AF003522 | 3162 bp mRNA PRI 17-JUN-1997 |
| LOCUS | DEFINITION | Homo sapiens Delta mRNA, complete cds. | |
| ACCESSION | AF003522 | | |
| NID | 92197068 | | |
| KEYWORDS | human. | | |
| SOURCE | ORGANISM | Homo sapiens | |
| | | Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; | |
| | | Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; | |
| | | Homo. | |
| REFERENCE | 1 | (bases 1 to 3162) | |
| AUTHORS | Mann,R.S., Gray,G.E., Henrique,D., Ish-Horowicz,D. and | | |
| | Artavanis-Tsakonas,S. | | |
| JOURNAL | Human Delta | | |
| REFERENCE | 2 | (bases 1 to 3162) | |
| AUTHORS | Mann,R.S., Gray,G.E., Henrique,D., Ish-Horowicz,D. and | | |
| | Artavanis-Tsakonas,S. | | |
| JOURNAL | Submitted (09-MAY-1997) Howard Hughes Medical Institute, Yale | | |
| | University, 295 Congress Ave, New Haven, CT 06510, USA | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..3162 | | |
| | /organism="Homo sapiens" | | |

| | | |
|------------|----------------------------------------------------------|--------------------------------------------------------------------|
| CDS | /db_xref="taxon:9606" | |
| | 323..2494 | |
| | /note="ligand for Notch receptor" | |
| | /codon_start=1 | |
| | /product="Delta" | |
| | /db_xref="PID:g2197069" | |
| | /translation="MGRSALALAVLSALLCOVSSGVFELKQEFVNNKGLGNRNC | |
| | CRGAGPPPCACRTFRVCLRHQASVSPPEPTTGSAVTPVIGVDSFSLPDDGGADS | |
| | AFNSPIRFPFTGFTWFTSLIIIEALHSDPDLATENPERLISRLATQRLHVTGWS | |
| | QDLHSSGRTDLKYSYFVDEHYEGGCVFPRDDAFGHFTCGERGKVPNGWKG | |
| | PYCTPPICLPGDEQGFCDKPEKCRVWGORYCDECIKIRYPCGLHGTQCPWOCNC | |
| | QEWGSLFCNODLNYCTHKKCKNGATCTNTGGSYTSCRPYGTGATCELGIDCDP | |
| | SPKNGGSCDTLNSYSCTCPPGFYKKICELSNMTCADPCPENGRCSDSPGGYSCR | |
| | CPVSGFNCERKIDICSSPSNGAKYDLGDAYLCRCQAGFSGRHCDNDVDCASS | |
| | PCANGTCDRVNDFSCPTPGYTGRNCSAPVSRCEHAPCHNGATCCHERGHYVCECA | |
| | RYGPGNCFDLPVAVVDTLEKLEGQGFPPFWAVCAVGLVLLMLLGCAGAV | |
| | VCVRLRLQKRRPADPCRETETMNNLANCQREKDISVSIIGATQIKNTNKKADFHGD | |
| | HSADKNGFKARYPAVDYNLVDLKGDDTAVRAHAKSRDTCQPGSSGGEKGTPTTLR | |
| | GGEASERKRPDSGCTSKDTKYQSVVISEKDECVIATEV" | |
| BASE COUNT | 705 a | 873 c 902 g 582 t |
| ORIGIN | | |
| | Query Match | 4.5%; Score 131; DB 22; Length 3162; |
| | Best Local Similarity | 61.8%; Pred. No. 4.31e-69; |
| | Matches | 514; Conservative 0; Mismatches 305; Indels 13; Gaps 12; |
| Db | 462 | GGGGCGGGGGCCACGCC--GTGGCGCTGCGCGGACCTTCTTCGCGTGTGCTCAAGCAC 520 |
| QY | 295 | GGAGCGACGGGAAGTGCCTGGCAGCTGCAAGACGGGTTTCGCGTCTGCCTAAAGCAC 354 |
| Db | 521 | TACAGGCGAGCGTGTCCCGGAGCGCCTGCACCTACGCGAGCGCCGTCACCCCGTG 580 |
| QY | 355 | TACAGGCGCACCATGACACACCTCCCGTGCACCTACGGGACGTGATCAGCCCAT 414 |
| Db | 581 | CTGGGCGTGCACCTCTTCACTGTGCCGAGCGGGGGGCG-CC-GACTCGCGG--TTCAGC 637 |
| QY | 415 | CTCGGCGAAGACTCGGTCAATCTGACCGAGCGCCAGCGCTTCCAGAACAGGGCTTCAG 474 |
| Db | 638 | AACCCCATCGGTTCCCTTTCGGCTTCACTGTGCGGGGACCTTCTCTCTGATATTGAA 697 |
| QY | 475 | AATCCCATCGAGTTCCCTTCTCGTTCTCATGGCGGGTACCTTCTCGCTGATCGTCGAG 534 |
| Db | 698 | GCTCTCCACACAGATTCTCTTGATGACCTCGCAACAGAAAACCCAGAAAGACTCATCAGC 757 |
| QY | 535 | GC-CTGG-CAT-GATAGCAACAATAGCGGAATGCGCGAACCACAAAGCTCCTCATCCAG 591 |
| Db | 758 | CGCTTGGCCACCCAGGACCTTGCAGGTGGGCGAGAGTGGTCCCAGGACCTGCACAGC 817 |
| QY | 592 | CGACTCTTGGTGAGCAGGTACTGGAGGTGTCTCTCCGAATGGAGACGAACAGTCGGAA 651 |
| Db | 818 | AGCGCGCGCAGGACCTCAAGTACTCTTACCGCTTCTGTGTGACGAACACTACTACGGA 877 |
| QY | 652 | TCGAGTACACGTCGCTGGAGTACGATTTCCGTGTCACTTCGATCTCAACTACTACGGA 711 |
| Db | 878 | GAGGGTGTCTCGTTTTTCTGCGCTCCCGGGAGAGTGCCTTTCGGCCACTTCACTGTGGG 937 |
| QY | 712 | TCCGGCTGTCCAAAGTCTCTCCCGCGCCGCGAGGATTCATTTGGACACTCGACTTGTCTG 771 |
| Db | 938 | GAGCGTGGGAGAAAGTGTGCAACCTGCTGCAAGGGGCCCTTACTGCACAGAGCCGATC 997 |
| QY | 772 | GAGACGGGCGAAATATATCTTTTACCGGATGGCAGGGCGATTACTGTCAATACCCAAA 831 |
| Db | 998 | TGCTTCCTTGGATGTGATGAGCAGCATGGATTTTGTGACAAACCCAGGGGAATGCAAGTGC 1057 |
| QY | 832 | TGC--GCCAA-A-G-CTGTGAA-CATGGACATTGGCACAACCCATCAATGCGTTTC 885 |
| Db | 1058 | AGAGTGGGTGCGAGCGCGGTACTGTGACAGTGTATCCGCTATCCAGGCTGTCTCCAT 1117 |
| QY | 886 | CAACTGGGTGGAAGGAGGCTTGTGCAACGAGTGTGCTTCTGGAACCCGAACATGCATCCAT 945 |
| Db | 1118 | GGCAGCTGCCAGGCGCTTGGCAGTGCACACTGCCAGGAAGCTGGGGGGGCTTTTCTGC 1177 |
| QY | 946 | GGCAGCTGCAACAAACCTTGACTTCATCTGCAACGAGGTTTGGGAGGCTGTGATGTC 1005 |

| | | | |
|----|------|---------------------------------------------------------------|------|
| Db | 1178 | AACCGAGGACCTAAGTACACGACACCAACCTAAGCCCTGCAAGAAAGGACCAACCGAC | 1237 |
| 0Y | 1006 | AACCGAGATCTAAGTACTACTGACCAACACACACCCCTGCAAGAAATGGCGGACCTGCTTC | 1065 |
| Db | 1238 | AACAGCGGCGGACGAGCTACACTTTCCTCCGCGCTGGGTACACAGTGTG | 1289 |
| 0Y | 1066 | AACACCGGCGAGGATTTGTACACATGCAATCGGCTCCAGGATACAGTGTGTG | 1117 |

RESULT8

LOCUS

DEFINITION

ACCESSION

NID

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

5'UTR

gene

CDS

sig_peptide

3'UTR

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

DRDELTAID

D.rerio mRNA for Deltad protein.

Y11760

g1888391

deltad gene.

zebrafish.

Danio rerio

Eukaryotes; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.

1 (bases 1 to 2342)

Dornseifer, P., Takke, C. and Campos-Ortega, J.A.

Overexpression of a zebrafish homologue of the Drosophila neurogenic gene Delta perturbs differentiation of primary neurons and somite development

Mech. Dev. 63 (2), 159-171 (1997)

97346722

2 (bases 1 to 2342)

Campos-Ortega, J.A.

Direct submission

Submitted (11-MAR-1997) J.A. Campos-Ortega, Institut fuer Entwicklungsbiologie der Universitaet zu Koeln, Gynofstr. 17, D-50923, Koeln, FRG

Location/Qualifiers

1. .2342

/organism="Danio rerio"

/db_xref="taxon:7955"

<1. .8

9. .2162

/gene="deltad"

9. .2162

/gene="deltad"

/codon_start=1

/product="deltad transmembrane protein"

/db_xref="PID:g1888392"

/db_xref="PID:g1888392"

/translation="MGRMLAVLLCVMISGFCGSEVFLKQELINKKGVGMANCKGSAEAGHCCECTFEFFRICLKHVQANVSPDPCPTGGATPVLAGSNSEFVPSFPDSSSLNPFPFARFTMPGTFSLIIDLHSDSDSDSTEDPDLISRMTORLITVEMESQILQVGRDELKTSYRFCDENHYEGECSEYCARPDRIHFHFGEGEELIICNSMGQVQCTEPLCGDCEGDHGFCKPCEKCRVGFSGKQCDCTIRYGCILHGTGTOOPMOCCODGWGICFMODLNTGCTHHRPCONGATCTNTGGGATGSCRPFTGDSCEIENVNFGSGSCRNGSCITDENYTSCTPPCFYGRNCELSAMTCADGCFNMGHCADNPBEGYTCQCCMGVGFNCEKIDHCSNPNCSNDACDLVDSYLQCEPFTGHCEDNIDECAITPQONGTGCDLSDYCTCPPGYGRNCSAVNKLCPNHCATCHEMDNRVCACTIPQYGRNCGFLPEPNOGQALVIGADKRTSYEEDDGGFPMTAGACGIIIVLVLIGGSYAVIYIRLKLOORSQOIDSHEISELMNLTNNSREKDLISIIIGATGVNINKKDFQFGDGNFGFSRSLVDYNIYVHLKQEDLGKEDSERSEARNCPLDSDSEKRNHLKSSSEKRTESLCKDKRTKYGSVFLSEKDECIATLEY"

18. .41

/gene="deltad"

2163. .>2342

629 a

590 c

606 g

517 t

4.0%; Score 115; DB 17; Length 2342;

Best Local Similarity 65.1%; Pred. No. 1.72e-57;

Matches 313; Conservative 0; Mismatches 162; Indels 6; Gaps 4;

512

GCTGAAGTACTATACAGATTCGTTTGATGATGACATTACGCGAGAGCGTGCTCGT

571

666

GCTGAGTACGATTCCTCGGTGATCCTGCAGATTCACATACATACGAGATCCGGCTGTGCCAA

725

| | | | |
|------------|------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|------|
| Db | 572 | CTTCCGCCGTCGGGCGCATGATCTATTTCGGGCACCTTCACCTGGGAGAGCGCGGGAAT | 631 |
| Oy | 726 | GTTCCTCCGCGCCCGCCGCGCATCTTTTGACACTTCGACTTGTCTGGAGACGGGCGGAAT | 785 |
| Db | 632 | TATCTGCACACTCCGGATGGAAGAGCAGTACTGCACAGAACCAATCTGTCTTCGCGGGTG | 691 |
| Oy | 786 | TATCTGTTTAGACCGGATGGAAGGCGGATTAATCTGATACATACCAAAATG-CCGCAAAAGC- | 843 |
| Db | 632 | TGATGAGACCATGAGCTTTTGGACAAACCCGGTGATGACAAATGCAGATGAGATTAG | 751 |
| Oy | 844 | TG-TGAA---CATGAGCACTTCGACAAACCAATCAATGCGTTTGCCCAACTGGGGGTGAA | 899 |
| Db | 752 | TGGAAGTACGTGTGACGCTGATCGCTTACCACGAGCGTGTGATGCTGACCTGCGCAACA | 811 |
| Oy | 900 | GGGAGCCTTGTGCACACGAGTGTGCTTGAAACCCAGCTGATCCATGGCACCTGGAACAA | 959 |
| Db | 812 | GCCCTGGCAATGCAACTGCCAAGAGGTTTGGGGAGGTCCTCTCTGTATTAACCAAGATCTCA | 871 |
| Oy | 960 | ACCGTGGACTTGTGCATCTGTGCACAGAGGGTTGGGGAGGCTGTATGACCAACAGGATCTGA | 1019 |
| Db | 872 | TTACTGCACACATCAACAAGCCGTGCGAGATGAGACCACTTGACCAACACAGGCCAGG | 931 |
| Oy | 1020 | CTACGCGACCAACCAACACACCTGCAAAATGCGCGGAACCTGCTTCAACACCGCGGAGG | 1079 |
| Db | 932 | AAGCTACACGTCGTCTACAGACCTGGGCTTACCGGGGAGACAGCTGTGATTTAGAGTCA | 991 |
| Oy | 1080 | ATTGTACACATGCAAAATCGCTCCAGATTAAGTGTATGTATGCGAAATGAGATCTA | 1139 |
| Db | 992 | C 992 | |
| Oy | 1140 | C 1140 | |
| RESULT | 9 | | |
| LOCUS | RNU78889 | 2795 bp | mRNA |
| DEFINITION | Rattus norvegicus Delta1 mRNA, complete cds. | | |
| ACCESSION | U78889 | | |
| NID | g1699045 | | |
| KEYWORDS | | | |
| SOURCE | Norway rat. | | |
| ORGANISM | Rattus norvegicus | | |
| REFERENCE | 1 (bases 1 to 2795) | | |
| AUTHORS | Disibio,G., Hebshl,L., Boulter,J. and Weimaster,G. | | |
| JOURNAL | CDNA sequence of R. norvegicus (rat) Delta1 | | |
| REFERENCE | 2 (bases 1 to 2795) | | |
| AUTHORS | Disibio,G., Hebshl,L., Boulter,J. and Weimaster,G. | | |
| JOURNAL | Direct Submission | | |
| FEATURES | Submitted (19-NOV-1996) Biological Chemistry, UCLA School of Medicine, 10833 Leconte Avenue, CHS 33-257, Los Angeles, CA 90095-1737, USA | | |
| source | location/Qualifiers | | |
| | 1. 2795 | | |
| | /organism="Rattus norvegicus" | | |
| | /db_xref="taxon:10116" | | |
| | 290..2434 | | |
| | /note="DSL ligand" | | |
| | /codon_start=1 | | |
| | /product="Delta1" | | |
| | /db_xref="PID:g1699046" | | |
| | /translation="MGRRSALALAVSALICQWSSGVFELKIOEFVNRKGLGNRNC | | |
| | CRGSGSPKACRFPFRYCLHYOASVSPPEPTYGSAVTAVLGYDSSLPGAIDPRA | | |
| | FSNPRTFPGFTWPGFTSLILEALHTDSPDLATENPERLISRTTTRKALTVGEMSO | | |
| | DLHSSGRTDLRYSRFVCDDEHYEGEGSVPCRPDDAFGHFTGEREKMDPEWKGO | | |
| | YCDPILCPGCDQDHGCRKPGDEKCGVNGRCDCEICIRYPGILHCTCOOPMCONCO | | |
| | EGMGGLKQNDLNYCTHHRKRNAGATNTNGOSYTGCSRPGYTGACIELEVCAPS | | |
| | PCNGSGCTDLEDSYCTCPGPGYGVKVCISAMTCDGCPFNNGRGSDNDGVTGHC | | |
| | PAPFSGNCKRKIDLCSSSCSNQAKVYVDLGNSTLCAQCGTGFSGRYCEDNVDDCASP | | |
| | CANAGCTRDSVNDSCTCPGTYGRNCSAVPSREBAFNGATCRORGORTMECAQ | | |

GYGANCQFLLPEPPDLIAAOGGSPWVAVCAGVVLVLLLLGCAAVVVRLLKLO
KQPPDPCCGGTETNNLANCQREKDVSIIGATQIKNTNKKADFGHDGADKSSF
KARYPTVDYNLRDLKQDEATVRDAHSKRDKTCQOSGSGVEEKSTSLRGGVEPDRKR
PESVYTDKTIQSVILSAEKDECVIATEV

BASE COUNT 632 a 788 c 806 g 569 t
ORIGIN

Query Match 3.8%; Score 111; DB 23; Length 2795;
Best Local Similarity 67.08; Pred. No. 1.29e-54;
Matches 278; Conservative 0; Mismatches 131; Indels 6; Gaps 4;

Db 831 ACTACATGAGGAAGTGTCTCGTGTCTGCGCCGAGCGGGATGATGCTTTGGCCACT 890
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 701 ACTACTACGGATCGGCTGTGCCAAGTCTGCCGCGCCGCGGAGTTCATTGGACACT 760
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 891 TCACCTGCGGGGAGAGAGGAGAGATGTCGACCTTGGTGAAAGGCGGACTGTCGA 950
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 761 CGACTTGTCTGGAGACGCGGGAATATCTGTTTACCGGATGCGAGGGCGATTACTGTC 820
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 951 CTGACCCCAATTGCTGCCAGGCTGTGATGACCAACATGATATGTCACAAACCGGGG 1010
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 821 ACATACCCCAATG-C-GCCAAAGGCTG-TGA---ACATGGACATTCGACAAACCCCAATC 874
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1011 AATGCAAGTCAGAGTGGCTGGCAGGCGCGCTACTGCGATGAATGATCCGATACCCAG 1070
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 875 AATGCGTTTCCCACTGGCTGGAGGAGGAGCTTGTGCAAGAGTGGCTTGGAAACGA 934
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1071 GCTGTCTCCATGTTACTGCGCAGAGCCCTGGCAGTGAATGCGCAAGGCTGGGGG 1130
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 935 ACTGATCCATGCACTGCAACAAACCTTGGACTTGCATCTGCAACGAGGTTGGGAG 994
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1131 GCTCTTCTGCAACAGGATCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 995 GCTTGTACTGCAACAGGATCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1191 CCACCTGCAACACAGGCGGAGGAGTACATGCTCTTGGCCGACCCGGTA 1245
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 1055 GAACCTGCTCAACACCGGAGGAGTGTACATGCAAAATGCTGCTCAGGATA 1109
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

RESULT 10 MWDELTA1 2857 bp RNA ROD 29-SEP-1995
LOCUS M.musculus mRNA for Delta-like 1 protein.
DEFINITION X80903
NID 9806569
KEYWORDS delta-like 1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2857)
AUTHORS Bettenhausen,B., de Angelis,M.H., Simon,D., Guenet,J.L. and Gossler,A.
TITLE Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta
JOURNAL Development 121 (8), 2407-2418 (1995)
MEDLINE 95401858
REFERENCE 2 (bases 1 to 2857)
AUTHORS Bettenhausen,B.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1994) B. Bettenhausen, Max-Delbrueck-Labor, Carl-von-Linnee-Weg 10, 50829 Koeln, FRG
REMARK revised by author 11-MAY-95
FEATURES Location/Qualifiers
1..2857
/organism="Mus musculus"
/strain="Balb/c x C57BL/6"
/db_xref="taxon:10090"
/dev_stage="d 10.5 p.c."
/clone_lib="lambda Zap II"
/chromosome="17"
/map="8.5 (EUCIB resource)"

gene
CDS

14..2182
/gene="Dll1"
14..2182
/gene="Dll1"
/codon_start=1
/product="DELTA-like 1"
/db_xref="PID:g806570"
/translation="MGRSALALAVSALLCOVSSGVFELKLOEVNKKLLGNRNC
CRGSGPPACRFTFFVCLKHYOASVSPPCPYGSNTPVLGVDSFSLPDGAGIDPA
FSPNRPFGTWPFTSLIEALHTDSPDLATENPERLSITLTHRHUTVGEWSQ
DLHSRGTDLRYSYRFVDEHYEGCVFCRPRDDAFHFTCGDERKMDPGWKGQ
YCDPILCPGDQHGCKRQVQWQRYCDECIYPCGLHGTQQPQWQNCQ
EGWGLFCNODLNYCTHHRPCRNATCTNTGQSYTSCSPRYTGANCELEVEDCAPS
PCKGASCTDLESFCTCPGPGVGVKVELSAMTCADGPCFNGRSDNPDDGVTCHC
PLGSGFNEKKMDLGGSPCSNGAKCYDLGNLSYLCRQAGFGRVCEDNVDCAASP
CANGCTRDSVNDFTSCTPPYTGKNCAPVSRCEHAPCHNGATCQHQGRYMWCECAQ
GYGPNCOFLPPPPGPMVLDLSERHMSQGGFPWAVAGCAVWLVLLLLGCAAVV
VCVRLKQHQPPPECGGETETMNLNQCQREKDVSVIIGATQIKNTNKKADFGHD
HGAKKSSFKVRYPTVDYNLRDLKQDEATVRDTHSKRDKTCQOSGSGVEEKIAPTLLRG
GEIPDRKRPESVYSTSKDTKYQSVYVLSAEKDECVIATEV"
29..61
sig_peptide
repeat_region
/gene="Dll1"
674..1564
/rpt_family="EGF-like"
/rpt_unit=674..769
/rpt_unit=770..871
/rpt_unit=872..991
/rpt_unit=992..1005
/rpt_unit=1106..1122
/rpt_unit=1123..1136
/rpt_unit=1337..1450
/rpt_unit=1451..1564
BASE COUNT 656 a 751 c 783 g 667 t
ORIGIN

Query Match 3.7%; Score 108; DB 23; Length 2857;
Best Local Similarity 61.6%; Pred. No. 1.81e-52;
Matches 498; Conservative 0; Mismatches 294; Indels 16; Gaps 13;

Db 175 CTGAGGACCTTCTTTGCGGTATGCTCAAGCACTACACAGCCAGCGTGTCCAGGAGCC 234
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 321 CTGCAAGCGGGTTTCGGGTCTGCTTAAAGCACTACACAGCCACCATCGACACACCTC 380
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 235 ACCCTGACCTACGCGAGTGTCTGACGCCAGTGTGGTGTGCTGCTCTTACGCTGCC 294
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 381 CCAGTGACCTACGCGGACGTGATCAGCCCATCTCTCGCGGAGAACTCGGTAATCTGAC 440
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 295 TGATGGCGCAGGCATC--GACCCGCG--CTTCAGCAACCCCATCCGATTCCTCCGCTT 351
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 441 CGAGCCCGCGCTTCCAGAAAGGGCTTCACGAATCCCATCCAGTTCCTCTCTCGTT 500
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 352 CACCTGCGCAGGTACCTTCTCTGTATCATTAAGACCCCTCCATACAGACTCTCCCGATGA 411
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 501 CTCATGCGCGGTACCTTCTGCTGATCTGCTGAGCCCTGGCATG-ATACGAACAATACGG 559
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 412 CCTCGCAACAGAAACCCAGAAAGACTATCATCGCCGCTGACACACAGAGACACTTAC 471
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 560 GCAATGCGC-GAAGCAACA-AGCTCTCATCCAGCACTCTTGGTGACAGGACTACTGGA 617
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 472 TGTGGGAGAAAGTGTCTCAGGACCT-TCACAGTAGCGCGCCGACACACCTCCGTTACT 530
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 618 GGTGTCTCCGAATGGAAGAACGAAGTTCGGAATCGCAGTA-CACGTGCTGGAGTACG 676
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 531 CTTACCGGTTTGTGTGTGACGAGCACTACTACGGAAGAGTGTGCTGTGTTCTGCCGAC 590
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 677 ATTTCGCTGTCACCTGCGATCTCACTACTACGATCGGCTGTCCCAAGTCTTCCGCGC 736
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 591 CTCGGGATGACGCTTTGGCCACTTCACTCTCGGGGACAGAGGGGAGAAAGATGTGCGAC 650
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 737 CCCGCGAGGATTCATTTGGACACTCGACTTGTCTCGAGAGCGGGGAAATATCTGTTTGA 796
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 651 CTGCTGGAAGGCGGACTGTCGCTGACCCCAATCTGCTGCCAGGGTGTGATGACCA 709
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

QY 797 CCGATGCGAGGCGGATGCTACTGTCACATACCCAAA-TG-C-CGCCAAGGCTG-TGA---A 849

Db 710 CAGGATACAGTGTGACAAACCGAGGAGTGCAGAGAGTGGCTGGCAGGCGCGCTAC 769

QY 850 CATTGACATGCGACAAACCAATCAATGCGTTGGCCAACTGGGCTGGAGGAGCCCTTG 909

Db 770 TGGCATGATGTCATCCGATACCCAGGTTGTCTCCATGGCAGCTGACACACCCCTGGCAG 829

QY 910 TGCACAGAGTGGCTTGTGGAACCGAATGCTGATGCTGACCTGCAACAAACCCCTGGACT 969

Db 830 TGTACTGCGAGAGAGCTGGGGGCGCTTTTGTGCAACCAAGACTGACTGACT 889

QY 970 TGCATCTGCAACGAGGTTGGGAGGCTTGTACTGCAACCAAGAGTGTGACTGACAC 1029

Db 890 CACATAGCCGTGAGAGAGAGCCCTGTCACCAACAGGCGGAGGAGAGTACAC 949

QY 1030 AACACAGACCCCTGCAAGAAAGGCGGAACCTGCTTCAACACCGGAGGAGTGTACACA 1089

Db 950 TGTCTCGCGAGCTGGGTATACAGTG 977

QY 1090 TGCAATGCGCTCCAGATACAGTGTG 1117

RESULT 11

LOCUS AF006488 3007 bp mRNA VRT 16-JAN-1998

DEFINITION Danio rerio deltab mRNA, complete cds.

ACCESSION AF006488

NID 92772824

KEYWORDS zebrafish.

SOURCE Danio rerio

ORGANISM Eukaryotes; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprininae; Rasbora; Danio.

REFERENCE 1 (bases 1 to 3007)

AUTHORS Haddon,C., Smithers,L., Schneider-Maunoury,S., Coche,T., Henrique,D. and Lewis,J.

TITLE Multiple delta genes and lateral inhibition in zebrafish primary neurogenesis

JOURNAL Development 125 (3), 359-370 (1998)

MEDLINE 98165391

REFERENCE 2 (bases 1 to 3007)

AUTHORS Haddon,C.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1997) Vertebrate Development Lab, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

FEATURES

source location/Qualifiers

1..3007

/organism="Danio rerio"

/db_xref="taxon:7955"

/dev_stage="20-28 hour"

/clone_lib="from Dr. D. Grunwald, made by R. Riggelman and K. Heide"

1..3007

/gene="deltab"

191..2038

/gene="deltab"

/note="transmembrane protein with 8 EGF-like repeats in its extracellular region; similar to Drosophila Delta protein, encoded by GenBank Accession Number Y00222, Mus musculus Delta-like 1 protein, encoded by GenBank Accession Number X80903, Xenopus laevis X-Delta-1 gene, encoded by GenBank Accession Number L42229, and Chick C-Delta-1 gene, encoded by GenBank Accession Number U26590"

/codon_start=1

/product="deltab"

/db_xref="PID:g2772825"

/translation="MAHLISLYCLSVSLQLVASGVFELKVSFTTRRRCRTDCCNIPFRICLSDSEVIAEPCTGTGOTSVLRADDSIASAARVDFHFKMGFTSLIIFANNAESPKEHDYIENONNLISRLATPRLAYGDSADVHFGDPSRLSYHFCDFEYTGACSDYCRPRDITLGHITCDENKRECLVGMODYCSDPLCSSDCSRBYCESPGECKRCLGWGSPSCSECVHPGLHGTCSQPMQVCVCEKMGGLFNMOLNYCTN

BASE COUNT 787 a 634 c 764 g 822 t

ORIGIN

Query Match 3.5%; Score 100; DB 17; Length 3007;

Best Local Similarity 65.8%; Pred. No. 8,92e-47;

Matches 271; Conservative 0; Mismatches 135; Indels 6; Gaps 3;

Db 683 TACTTGGGCGAGGCTGTGTCAGATTCAGCGCTCCGTCGAGCAGCAGCTGCGCCACTAT 742

QY 703 TACTAGGATCCGCTGCTGCCAAGTTCTGCGGCGCCGAGAGATTCATTTGGACACTG 762

Db 743 ACTTCGATGAGAAATGAGACAGCAAGATGTTGTGGATGCGAGGAGACTGCTCT 802

QY 763 ACTTCTCGAGAGAGCGGGAATATTCGTTGACCGGATGCGAGGCGGATTAAGTCTGAC 822

Db 803 GACCCATCTGCTCTTCACTGACATGAGACAGATGTTATGCGAGTCTCCGAGAG 862

QY 823 ATACCAATGCGC--CAAG-GCTGTGAC--ATGACATTTGCGACAAACCCCAATCAA 876

Db 863 TGCAGTGTGCTGTGGGGGCGAGGCGCTCCGTCGAGGAGTGGCTCATTTATCAGGA 922

QY 877 TCGTTTCCCAACTGGGCTGGAGGAGGAGCTTGTCAACAGAGTGGCTTGTGACAGCAG 936

Db 923 TGCCATGAGGAGCTGTTCAACAGCAGCATGAGTGTGTGAAGAGAGGTTGGGAGAGC 982

QY 937 TGCATCAGTGGCACTGCAACAGCAGCTGCACTGCACTGCAACAGAGGTTGGGAGAGC 996

Db 983 CTCTTCTGTAACAGAGACCTCACTGTAACCAATCAACAAACCTGCGCTTAAGCGCC 1042

QY 997 TTGTACTCAACAGAGATCTGACTGACCAACCAACAGACCTGCAACAAATGCGCGA 1056

Db 1043 ACCGCAACCAACAGAGGAGGCGAGCTGACACTGCGGCGCGGAGAT 1094

QY 1057 ACTGCTTCAACCGCGGAGGAGTGTGACATGCAATGCGCTCCAGAGAT 1108

RESULT 12

LOCUS XELXDEL 2166 bp DNA VRT 17-JUL-1995

DEFINITION Xenopus laevis X-Delta-1 gene, complete cds.

ACCESSION L42229

NID 9807695

KEYWORDS Delta; Notch; X-Delta-1; neurogenic gene.

SOURCE Xenopus laevis DNA.

ORGANISM Eukaryotes; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 2166)

AUTHORS Henrique,D., Adam,J., Myat,A., Chitnis,A., Lewis,J. and Ish-Horowitz,D.

TITLE Expression of a Delta homologue in prospective neurons in the chick

JOURNAL Nature 375 (6534), 787-790 (1995)

MEDLINE 95319507

REFERENCE 2 (bases 1 to 2166)

AUTHORS Chitnis,A., Henrique,D., Lewis,J., Ish-Horowitz,D. and Kintner,C.

TITLE Primary neurogenesis in Xenopus embryos regulated by a homologue of the Drosophila neurogenic gene Delta

JOURNAL Nature 375 (6534), 761-766 (1995)

MEDLINE 95319503

REFERENCE 3 (bases 1 to 2166)

AUTHORS Ish-Horowitz,D.

TITLE Direct Submission

JOURNAL Submitted (15-MAY-1995) David Ish-Horowitz, Developmental Biology Unit, ICRF, Oxford, UK

FEATURES

source location/Qualifiers

1..2166

/organism="Xenopus laevis"

```
/note="(vector lambda gt10)"
/db_xref="taxon:8355"
/clone_lib="Chris Kintner's St17"
1..2166
CDS
/product="x-Delta-1"
/db_xref="PID:g807696"
/translation="MGQRMFLTLVLSAVLCQISCSGLFELRQEFVNVKGLLGNMNC
CRPLASIQKECTFRICLKHQSNVPEPCTYGAVTPVLGTSFVVPSSNA
DPTESNPRFPFGFTWPGTFLIIIEAHDSADDLNTENPERLISRLATQRLHLYGEQ
WSQDHSRDELKYSYFVCEYDYYGEGSCDYCRPDADAFHFGSCGKGLKCPGW
KGLYEIPCLPCDDEHGYCDKPGCEKRCVWGQYDECIRYPCGLHGTQQQWQC
NCOEWMGLFCNLDNYCTHHKPCENGATCTNTGOSYTCSCRPYVTSNCEIYNEC
DANCPKNGSSDLNSYTCSPGFGKNCELSANTCADGCPNGRCADNDGQYI
CFPCVYSGFNEKKIDYCSNPANGARCELDGNSYICQCEGSGRNCDDNDQCT
SFPQGVGCEQNDINDYCTPCPGYIGKNSMPTIKCHNPCHNGATCERNRYVCO
CARGTGNNGCQFLLEPVPVVDLTERYEGOSQGPFWIAYGACGLVLMLLGCAAV
VVCVRVYQKRHRQPEACRGSKTWNLANCOREKDISVFTGTQIKNTNKKIDFLS
ESNNEKNGYKPRYPSVDYNLHVKLNEDSPKERSKCAKSSNDSDESDVNSVHSKR
DSSERRRPSAYSTSKDTKYQSVVYISDEKDCIATEV"
BASE COUNT 586 a 514 c 541 g 525 t
ORIGIN
Query Match 3.3%; Score 94; DB 17; Length 2166;
Best Local Similarity 67.7%; Pred. No. 1.53e-42;
Matches 180; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Db 706 CATGATATGTGACAAGCCTGGGAATGCAATGATGAGTGGGTGGCAAGCCGCTAC 765
QY 850 CATGACATGTGCAAAACCAATCAATCGTTTGGCAACTGGGTGGGAAGGAGCCTTG 909
Db 766 TGTGATGATGATCTGTTACCCAGATGCTGATGATGATGATGATGATGATGATG 825
QY 910 TGAACGAGTGGCTTGTGAACGGAATGATGATGATGATGATGATGATGATGATG 969
Db 826 TGCAACTGCAAGAAGCCTGGGAGGACTTTCTGTAACCAAGATCTTAACTACTGACC 885
QY 970 TGCATCTGCAAGAGGCTTGGGAGGCTTGTACTGCAACCAAGATCTGACTACTGACC 1029
Db 886 CACCACAAGCCTGCAAAATGAGGACCATGCAACCAAGCCTGCAACCAAGCCTGCAAC 945
QY 1030 AACCACAGACCTTGAAGAATGCGGAACCTGCTTCAACACCGGCGAGGATGTACACA 1089
Db 946 TGCTCTGTGCTGCTGGGTACTGCTGG 971
QY 1090 TGAATGCGTCCAGGATACATGG 1115
RESULT 13
LOCUS AF030031 3051 bp mRNA VRT 27-JAN-1998
DEFINITION Danio rerio DeltaA (deltaA) mRNA, complete cds.
ACCESSION AF030031
NID g2809388
KEYWORDS zebrafish.
SOURCE Danio rerio
ORGANISM Danio rerio
Eukaryote; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinoidae; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 3051)
AUTHORS Appel,B. and Eisen,J.S.
TITLE Regulation of neuronal specification in the zebrafish spinal cord
by Delta function
JOURNAL Development 125 (3), 371-380 (1998)
MEDLINE 98165392
REFERENCE 2 (bases 1 to 3051)
AUTHORS Appel,B. and Eisen,J.S.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Institute of Neuroscience, University of
Oregon, 1254 University of Oregon, Eugene, OR 97403, USA
FEATURES
LOCATION/Qualifiers
1..3051
/organism="Danio rerio"
```

```
/db_xref="taxon:7955"
/chromosome="LG I"
/map="between 21173 and 2450"
1..3051
CDS
/genes="deltaA"
/notes="similar to Drosophila melanogaster Delta"
/codon_start=1
/product="DeltaA"
/db_xref="PID:g2809389"
/translation="MNQAKREAFNTROPFLDLSAKLRLDFHKLFIYLFIMGRHLL
LILFSLMLCOASSGVFLKLOELFKNGKQKNGKCGGLTTSYQOCCKTFE
ICLKHQNASPEPCTYGGTVPVLSNSFOVPTLPDGSFNTIRNFGTWPQTF
SLIIEALHADSDELITENPERLISRLATQRLHLYGEQSDQLHVSQTELYSRFV
DEHYEGECVFCRDPDADAFHFGSCGKGLKCPGWQYDECIRYPCGLHGTQQQWQC
CEKPGCEKRCVWGQYDECIRYPCGLHGTQQQWQCCEKPGCEKRCVWGQYDECIRY
TCPPGFGKNCELSANTCADGCPNGRCADNDGQYFCQCPYVTSNCEIYNEC
SSPNSGARCVLNSYLCOCPDGFTGMNCDRAGDECMSYPCQNGTQCEGAGTMC
CPGYTGRCSSPVSRCQHPNCHNGATCERNRYVCAVSYGSGRNCQFLPDRASQ
IASDVPWAVGSGVLVLLVAVACVAVVVCVRSKVRQRRRDEDEVANGENETINLT
NNCHRDKLAVSVGVAPVKNKIDFSSDDHLLSTTEKRSYKTRAPADYNLVHE
VKFEVKEVKEHAGKETMANELSDSCEDIKCSLQSDSECTEKKRRRLKSDASEK
SKYSESRYSESSESKYSESKYSCYSCYSCYSCYSCYSCYSCYSCYSCYSCYSCY
GGRK"
BASE COUNT 798 a 722 c 808 g 723 t
ORIGIN
Query Match 2.7%; Score 79; DB 17; Length 3051;
Best Local Similarity 62.1%; Pred. No. 3.92e-32;
Matches 279; Conservative 0; Mismatches 164; Indels 6; Gaps 4;
Db 916 GCTCAAGTACTCTACCGTTTGTGTGGCAGCAACTACTATGTTGGAAGGATGCTCAGT 975
QY 666 GCTGGAGTACGATTCGCTGTCACCTGCGATCTCAACTACTAGGATCCGCTGTCGCA 725
Db 976 GTTCTCGCGCCCGCAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1035
QY 726 GTTCTCGCGCCCGCAGAGATGATGATGATGATGATGATGATGATGATGATG 785
Db 1036 CATCTCGATGCTGGGTGGAAAGGCCAATACTGTCAGCAACCACTATCTGCTGCGAGG 1095
QY 786 TATCTGTTGACCGGATGCGAGGCGGATGATGATGATGATGATGATGATGATG 842
Db 1096 CGACGAGGACGAGATCTGTGAGAAACCTGCGAGTGTGAGTGTGAGTGTGAGTGT 1155
QY 843 CTGTGA--A-CATGGACATTCGCAACCAACCAATCAATGCTTTCGCAACCTGGCTG 899
Db 1156 AGGCGCTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1215
QY 900 GGGAGCCTTGTGCAACGAGTGCCTTGTGCAACCGAATGATGATGATGATGATGATG 959
Db 1216 GCGATGCAATGCACTGCGAGGAGGCTGGGTGGGTGCTTCTGCAACCAAGATCTGAA 1275
QY 960 ACCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019
Db 1276 CTACTGACACATGATAAGCCCTGCTGTAATGAGGACCATTTGTAGTAACACGCGCCAG 1335
QY 1020 CTACTGCAACCAACCAACGACCTGCAAGATGCGGGAACCTGCTTCAACACCGCGG 1079
Db 1336 CAGCTACACCTGTCTGTGCGCCAGAT 1364
QY 1080 ATTGTACACATGCAAAATGCGCTCCAGAT 1108
RESULT 14
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID g2724471
KEYWORDS Unknown.
SOURCE
```


9757: .865 avg

CC and adhesive molecules, and study of its expression are possible using
CC the DNA and antibodies raised against the Notch and Delta proteins.
SQ Sequence 2892 BP; 739 A; 763 C; 808 G; 582 T;

Query Match 99.9%; Score 2888; DB 5; Length 2892;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2890; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|-----|-----------------------------------------------------------------|------|
| Db | 1 | gaaticggagggaattattcaaaacataaaacacataaaacaaatttgatgttgcgcaca | 60 |
| Qy | 1 | GAATCGGAGGAATTATTCAAACATAAACAATAAACAATTTGAGTAGTTGCCGCACA | 60 |
| Db | 61 | caacacacacacagcccggtgattattacactaaagcgacactcaatcccaaaaatca | 120 |
| Qy | 61 | CACACACACACAGCCCGTGGATTATTACTAAAGCGACACTCAATCCAAAATCA | 120 |
| Db | 121 | gcaacaaaacatcaataaacatgcattgattataatgtttattaaacagattcattgc | 180 |
| Qy | 121 | GCAACAAAACATCAATAAACATGATTTGATTAATGTTTATTAACAGATTTCATTGC | 180 |
| Db | 181 | ttcacagtcatcgtgcaggttcacagttccggcagctttgagttgcctgaagtcttc | 240 |
| Qy | 181 | TTCACAGTTCATGTCGAGGTTACAGTTCGGCAGCTTTGAGTTGGCGCTGAAGTACTTC | 240 |
| Db | 241 | acgaacgatacagggcggaacagaggtcgtcgtgcagcgggagtcggagcgagcg | 300 |
| Qy | 241 | ACGAACGATACAGGGCGGACAAACAGAGGTCGTCTGCAGCGGGAGTCGGACGGCG | 300 |
| Db | 301 | acgggcgaagtgcctggcagtcgaagacggttttcggtcgtcgttaagcactaccag | 360 |
| Qy | 301 | ACGGCAAGTGCCTGGCAGCTGCAAGACGGGTTTCGGCTCTGCCTAAAGCACTACCAG | 360 |
| Db | 361 | gccacatgcacacacccctccagtcacactacgagggagtcagtcacgccttcggc | 420 |
| Qy | 361 | GCCACCATGCACACACCTCCAGTCACCTACGGGAGCTGATCATCGCCATTCTCGGC | 420 |
| Db | 421 | gagaaactcgtcaactgcagcagccagcgttcctccagaacagaggttcacgaatccc | 480 |
| Qy | 421 | GAGAACTCGGTCAANTGTGACCGAGCGCCAGCGCTTCAGAACAGGGGCTTCAGAAATCCC | 480 |
| Db | 481 | atccagttccctctcgttctcatggcggttaccttctcgtgatcgtcggagcgcttg | 540 |
| Qy | 481 | ATCCAGTTCCCTCTCTCGTTCATGGCCGGTACCTTCTCGTGATCGTCGGAGCGCTGG | 540 |
| Db | 541 | catgatcacgaataagcggcaatgcgcgaaccaaagaagtcctcctccagcactcttg | 600 |
| Qy | 541 | CATGATACGAACAATAAGCGCAATGCGGAACCAACAAGCTCCTCATCCAGCGACTCTTG | 600 |
| Db | 601 | gtgcagcaggtactgaggtgtcctccgaatggaagacgaacaaagtgcggaatcgagtac | 660 |
| Qy | 601 | GTGCAGCAGGTACTGGAGGTGTCCTCCGAATGGAAGACGAACAAAGTCGGAATCGCAGTAC | 660 |
| Db | 661 | acgtcgtgaggtacgatcttcgtcactgcgactcactactacgactcggcgtgt | 720 |
| Qy | 661 | ACGTGCTGGAGTACGATTTCCGTGTCCACTGGGATCTCACTACTACGAGTCCGGCTGT | 720 |
| Db | 721 | gccaaattctgcggcccgccgacgatttcttggaacactcgcacttgcctggagacgggc | 780 |
| Qy | 721 | GCCAAATTCTGCGGCCCGCCGACGATTATTGGACACTCGACTTCTCGGAGACGGGC | 780 |
| Db | 781 | gaaattatctgtttgaccggatggcagggcgattactgtcacatacccaaatgcgcacaa | 840 |
| Qy | 781 | GAAATTATCTGTTGACCGGATGGCAGGGCGATTACTGTACATACCAATTCGCCCAAA | 840 |
| Db | 841 | ggctgtgaacatggacattgcgaacacgaatcaatgctttgccaactggcgtggaag | 900 |
| Qy | 841 | GGCTGTGAACATGGACATTGCGAACACCAACCAATCAATGCTTTGCCAACCTGGGCTGGAAG | 900 |
| Db | 901 | ggagccttgcgaacgagtcgttctggaacccgaactgcctccatggcactgcacacaa | 960 |
| Qy | 901 | GGAGCCTTGTGCAACGAGTCGTTCTGGAAACCGAAGTTCATGCACTGCACTGCAACAAA | 960 |
| Db | 961 | ccctggacttgcatacgaacaggggttggggaggttgccttgcatacgaatggcagtggggtg | 1020 |

| | | | |
|----|------|----------------------------------------------------------------|------|
| Qy | 961 | CCCTGACATTGCAATCTGCAACGAGGGTTGGGAGGCTTGTACTGCAACACAGGATCTGAAC | 1020 |
| Db | 1021 | tactgcaccaaacacagaccctgcaagaatggcggaacctgtctcaacacccggcgagga | 1080 |
| Qy | 1021 | TACTGCACCACACAGACCCCTGCAAGAATGGCGGAACCTGCTTCAACACCGCGGAGGA | 1080 |
| Db | 1081 | ttgtacacatgcaaatgcgtccagagatacagtggtgtagattgcgaaaaatgagatctac | 1140 |
| Qy | 1081 | TTGTACACATGCAAAATGCGCTCCAGGATACAGTGGTGATGATTCGCAAAATGAGATCTAC | 1140 |
| Db | 1141 | tcctgcgatccgcatcattccctcccaaatggtgggtacctgcactcagatgagcgcac | 1200 |
| Qy | 1141 | TCTGTGATGCCCATGTCAATCTCCAGAAATGGTGTACCTGCAATGCAATGAGAGCGCAC | 1200 |
| Db | 1201 | acaaaacccggtcacagtgatcattgcgcaacggctggagcggaagatgtgcgaggag | 1260 |
| Qy | 1201 | ACAAAACCCGGCTACAAGTGTCTATTCGCCCAACGGCTGGAGCGGAAAGATGTGCGAGGAG | 1260 |
| Db | 1261 | aaagtgcacgtgttctggacaaaacctgtcatcaggggaatctgcgcgcaacgttcgctct | 1320 |
| Qy | 1261 | AAAGTGTCTACGTGTTCGGACAAACCTGTCTATCAGGGAATCTGCCCAACGTTCGTCTCT | 1320 |
| Db | 1321 | ggcttgggaagcaaggtcaggtcaccagtcgaaatgtcccatgtgctctacagcgaccc | 1380 |
| Qy | 1321 | GGCTTGGGAAGCAAGGTCAGGCTACCACTGCGAATGTCCCATTTGCTACACGGGACCC | 1380 |
| Db | 1381 | aactgcgatctccagctggcaactgcagtcgccatccatcgcataaaacggtggaagtgt | 1440 |
| Qy | 1381 | AACTGCGATCTCCAGCTGGCAACTGCAGTCCGAATCCATGTCATAAACGGTGAAGCTGT | 1440 |
| Db | 1441 | cagccgagcggaagtgtatttcccagcggtatttccgggaacagagatcgagaccaac | 1500 |
| Qy | 1441 | CAGCCGAGCGGAAGTGTATTTCCCAAGCGGGATTTTCGGGAACGAGATCGGAGACCAAC | 1500 |
| Db | 1501 | attgcagattgtcttgccaccagtcgagacgagcgacccctgcataagatatggtcaac | 1560 |
| Qy | 1501 | ATTGCAATTGTCTTGCCACCACTGAGGACGAGACGAGGACCTGTCATAGATATGTTCAAC | 1560 |
| Db | 1561 | caatatcgtcgaatgcgttcccggtttccatggcaacccactgtagtagcaagtgtac | 1620 |
| Qy | 1561 | CAATATCGTGCCTAATGCGTTCCCGTTCCATGGCACCCACTGTAGTAGCAAAAGTTGAC | 1620 |
| Db | 1621 | ttgtgcctcatcagaccgtgtgccaatggagggaacctgtgtgaatctcaacaacgattac | 1680 |
| Qy | 1621 | TTGTGCTCATACAGCGTGTCCCAATGAGGAACCTGCTTGAATCTCAACAACGATTTAC | 1680 |
| Db | 1681 | cagtgcacctgtcgtcggaatttactggaagattgctctgtggacatcgatgagtg | 1740 |
| Qy | 1681 | CAGTGCACCTGTCTGTCGGGATTTACTTGGCAAGGATTTGCTCTGTGGACATCGATGAGTGC | 1740 |
| Db | 1741 | agcagtggaccctgtcataacggcggttcgatgaacccgctcaattcgttcgaatgc | 1800 |
| Qy | 1741 | AGCAGTGGACCTGTCTATAACGGCGGCACTTGCATGAACCGCGTCAATTCTGTCGAATGC | 1800 |
| Db | 1801 | gtgtgtcccaatggtttcaggggcaagcagtcgagtagagagtcctcagattcggtagcc | 1860 |
| Qy | 1801 | GTGTGTCCCAATGTTTTCAGGGGCAAGCAGTCGATGAGGAGTCTACGATTCGGTGACC | 1860 |
| Db | 1861 | ttcgatcccaacaataatgagcgaccacacaacagcgagagccatggtttgaccaatgcc | 1920 |
| Qy | 1861 | TTTCGATGCCCAACCAATATGGAGCGACCAACAACGAGAGCGGATGTTTGAACCAATGCC | 1920 |
| Db | 1921 | caggtagtcctaaattgctgttttccgtttgcgagtcgctttggcggttattgcggcg | 1980 |
| Qy | 1921 | CAGGTAGTCTAATTGCTGTTTCTCGTTCGATGCTTTCGTTGGTGGCGGTATTTCGCGCG | 1980 |
| Db | 1981 | tgctgtgtcttctgcataagaacgcaagcgttaagcgtctcaggaaaaagacgacgag | 2040 |
| Qy | 1981 | TGCTGTGTTCTCTGCATGAAGCGCAAGCGTAAGCGTGTCTCAGGAAAAAGGACGACGCGAG | 2040 |
| Db | 2041 | gcaaggagcagaacgacgaatgcggtggccaatgcatacaaatggcagtggggtg | 2100 |

DT 28-APR-1997 (first entry)
 DE C-Delta-1 gene (alternatively spliced variant).
 KW C-Delta-1; cell proliferation; nervous system disorder;
 KW tissue regeneration; Notch; cervix cancer; breast cancer;
 KW lung cancer; colon cancer; melanoma; seminoma;
 KW neurogenesis; therapy; ss.
 OS Gallus sp.
 FH Key Location/Qualifiers
 FT cds 277..2499
 FT /*tag= a
 PN WO9701571-A1.
 PD 16-JAN-1997.
 PF 28-JUN-1996; U11178.
 PR 28-JUN-1995; US-000589.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PA (UYIA) UNIV YALE.
 PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
 PI Lewis J;
 DR WPI: 97-100159/09.
 DR P-PSDB: W00876.
 DR New vertebrate Delta protein, DNA and antibodies - for treating and
 PT preventing cancer, nervous system disorders and for tissue
 PT regeneration
 PS Disclosure: Fig 1B1-1B2; 135pp; English.
 CC The C-delta-1 gene (T58897) codes for the chick homologue (W00876)
 CC of Drosophila Delta, a protein that binds to Notch protein. It was
 CC obtd. by PCR amplification of cDNA from stage 4-6 embryos using
 CC primers (see also T59455-56) based on fly Delta sequences. A
 CC shorter variant (T58897) was also identified. C-Delta-1 expression
 CC appears to be the earliest known marker for prospective neurons.
 CC Mouse (T58899) and human (T58900, T59454) Delta-1 sequences have
 CC also been isolated. Delta-1 genes can be used in the prodn. of
 CC Delta polypeptides and (including antisense sequences) utilised in
 CC the treatment of disorders of cell fate or differentiation, such as
 CC cancer, and nervous system disorders, or to promote tissue
 CC regeneration and repair.
 SQ Sequence 2883 BP; 638 A; 858 C; 792 G; 593 T;

Query Match 5.9%; Score 171; DB 28; Length 2883;
 Best Local Similarity 61.1%; Pred. No. 1.49e-96;
 Matches 590; Conservative 0; Mismatches 367; Indels 9; Gaps 7;

Db 334 caggttgacggtccgggtgttcgactgaagctgagagtggttgcacaagaagggg 393
 Qy 196 CAGGTTACAGTTCGGGACGCTTGAGTTGGCTGAGTTCAGCAACGATCACGGG 255
 Db 394 ctgctcagcaacgcgaactgtcgccggggggcgcccgagggcgccgagcagcag 453
 Qy 256 CGGGACAACGAGGGTCTCTCTCGAGCGGGGAGTCCGAGCGGAGCGGCAAGTGCCTG 315
 Db 454 tgcgactgcaagaccttcttcgctgctgctgaagcactaccagggcgctctcccc 513
 Qy 316 GGCAGCTGCAAGACGCGGTTTCGGCTGCTGCTTAAGCACTACCAGGCGACCATCGACACC 375
 Db 514 gagccgcttgacactagcagcagcgcacacccccctctcggcgccaactcttcagc 573
 Qy 376 ACTCCAGTGCACCTACGGGGAGCGTATCACGCCCTTCTCGCGGAGAACTCGGTCAAT 435
 Db 574 gtcccgacggcgggcgccgcagcccgcttcagcaaccccttcgcttcccccttc 633
 Qy 436 CTGACCGACGCCAGCGCTTCAGAACAAAGGGCTTCAGTAATCCATCCAGTTCCTTC 495
 Db 634 ggttcacctggcccgccacattctctgctcatcatcagaggtcttgacacagcactcccc 693
 Qy 496 TCGTTCTCTGCGGGGTACTCTTCTCGTGTATCGTCGAGGC-CTGG-CAT-GATACGAAC 552
 Db 694 gagacactaccacagaacccccgagcgcctcatatagcgccttgccaccacagagcac 753
 Qy 553 AATAGGGGCAATGCGGCAACCAACAACTCTCTCATCCAGCGACTCTTTGGTGCAGCAGTA 612
 Db 754 ctgaggtggcgagagtggtccagagacctgcacagcagcgcccgaccagacctcaag 813
 Qy 613 CTGAGGAGTGTCTCCGAATGGAGCAACCAAGTGGGAATCGGAGTACACGTCGCTGGAG 672

Db 814 tactctactcgcttinnngtggatgagcactactacaggggaagagctgctgtctcttcg 873
 Qy 673 TACGATTTCGGTGTACACTCGGATCTCAACTACGATACGATACCGGCTGTGCGCAAGTCTTCG 732
 Db 874 cggcccggtgacgacgcgtcttcggtcacttcacctgtgtgagagcgtggcgagaaggtctgc 933
 Qy 733 CGGCCCCGCGACGATTTCATTGGACACTCGACTTGTCTCGGAGACGGGGGAAATTTATCTGT 792
 Db 934 aaccaggtctggaaggggagccagtagtgcactgagccgattgttgcctgggtgtgacgag 993
 Qy 793 TTGACCGGATGGCAGGCGGATTTACTGTACATACCCAAATCGCCAAAGGCTGTGA--A- 849
 Db 994 cagcagcgtcttcgcaacacacctgggggaatgcaagtgcagagtgggttggcagggcg 1053
 Qy 850 CATG--GACAT-TGCGACAAACCAATCAATCAATGCGTTTGCCCACTGGGCTGGAAGGGAGCC 906
 Db 1054 tactgtgacgagtgatcccgatataccaggtgcctgcacgtgactctgcacgagccatgg 1113
 Qy 907 TTGTGCAACGAGTGGTCTCGAACCGAAGTGCATCCATGCACCTGCACAACACCTGG 966
 Db 1114 cagtgcactgcccaggaaggtggggcgcccttttctgcacaccagggacctgaactatgc 1173
 Qy 967 ACTTGCATCTGCAACGAGGGTGGGGAGGCTGTACTGCAACGAGGATCTGAACCTACTGC 1026
 Db 1174 actcaccacaagcatgcaagaatgggtgcacatgcacatgcacacacacccggtcagggagctac 1233
 Qy 1027 ACCAACACACAGACCTCGAAGATGGCGGAACCTGCTTCAACACCGCGGAGGATGTGTAC 1086
 Db 1234 actgtttcttcgacctggtgacacaggtccagctccagctgcagattgaaatcaacgaatg 1293
 Qy 1087 ACATGCAATGCGCTCCAGGATACAGTGGTGTATGCGGAATGAGATCTACTCTCTGC 1146
 Db 1294 gatgcc 1299
 Qy 1147 GATGCC 1152

RESULT 4
 ID T58897 standard; cDNA; 2088 BP.
 AC T58897;
 DT 28-APR-1997 (first entry)
 DE C-Delta-1 gene;
 KW C-Delta-1; cell proliferation; nervous system disorder;
 KW tissue regeneration; Notch; cervix cancer; breast cancer;
 KW lung cancer; colon cancer; melanoma; seminoma;
 KW neurogenesis; therapy; ss.
 OS Gallus sp.
 FH Key Location/Qualifiers
 FT cds 277..2463
 FT /*tag= a
 FT /note= "bases 1201-1260 do not encode the
 FT translated sequence for C-Delta-1
 FT polypeptide given in Fig 2, and
 FT bases 1981-2400 are missing from the
 FT sequence"
 PN WO9701571-A1.
 PD 16-JAN-1997.
 PF 28-JUN-1996; U11178.
 PR 28-JUN-1995; US-000589.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PA (UYIA) UNIV YALE.
 PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
 PI Lewis J;
 DR WPI: 97-100159/09.
 DR P-PSDB: W11719.
 DR New vertebrate Delta protein, DNA and antibodies - for treating and
 PT preventing cancer, nervous system disorders and for tissue
 PT regeneration
 PS Disclosure: Fig 1A1-1A3; 135pp; English.
 CC The C-delta-1 gene (T58897) codes for the chick homologue (W11719)
 CC of Drosophila Delta, a protein that binds to Notch protein. It was
 CC obtd. by PCR amplification of cDNA from stage 4-6 embryos using

CC primers (see also T59455-56) based on fly Delta sequences. A longer
CC variant (T58899) was also identified. C-Delta-1 expression appears
CC to be the earliest known marker for prospective neurons. Mouse
CC (T58899) and human (T59454) Delta-1 sequences have also
CC been isolated. Delta-1 genes can be used in the prodn. of Delta
CC polypeptides and (including antisense sequences) utilised in the
CC treatment of disorders of cell fate or differentiation, such as
CC cancer, and nervous system disorders, or to promote tissue
CC regeneration and repair.
SQ Sequence 2088 BP; 412 A; 664 C; 604 G; 408 T;

Query Match 5.1%; Score 147; DB 28; Length 2088;

Best Local Similarity 61.1%; Pred. No. 1,05e-79;
Matches 529; Conservative 0; Mismatches 328; Indels 9; Gaps 7;

DB 334 cagggttgcagctccgaggtgttcgagctgaagctgacagagttgtcaacaagaagg 393
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 196 CAGGTTGACAGTTCCGCGAGCTTGAGTGGCTGAAGTACTTCAGACAGATCAGCGG 255
DB 394 ctgtctcagcaacgcaactgtctcggaggggagcccgagagcgccggcagcagcag 453
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 256 CGGACACACAGAGGCTCCGCTGCGAGGAGTCCGACGACGCGCAAGTCCCTG 315
DB 454 tgcgactgcgaagactcttcctccggtctcgtgagacatacagagcgagctccccc 513
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 316 GCGAGCTGCAAGACGCGGTTCCGCTGCTTAAGACATCCAGGCCATCGACACC 375
DB 514 gaagccctgcagctacatacagcagccatcacccccctcctgcggcgaactcttcagc 573
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 376 ACCTCCAGTACAGCTACCTACGAGGAGTATGACGCCATTCCTGGGAGCAACTCGGTCAAT 435
DB 574 gtcccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 633
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 436 CTACACACACGCGAGCTTCCAGACAAAGGCTTCCAGAAATCCATCCATCCCTCCCTTC 495
DB 634 ggtctcagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 693
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 496 TCGTTTCAATGCGCGGCTACCTTCCGCTGATGCTGAGGCG-CTGGG-CAT-GATACGAC 552
DB 694 gacgacctcacacagaaacccccagcgagcgagcgagcgagcgagcgagcgagcgag 753
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 553 AATAGGCGCAATGCGGCAACCAACCAAGCTCTCATCCAGGACCTTGATGACAGGTA 612
DB 754 ctggcggtgagcgagagtggtccagagcgagcgagcgagcgagcgagcgagcgagcgag 813
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 613 CTGAGGCTGCTCTCGAATGGAAGCAAGTCCGAAATCGAGTACCTCCCTGAG 672
DB 814 tactctatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 873
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 673 TACGATTCGCTGACCTCGATCTCACTACTACGATCCGCTGCGCAAGTTCTGC 732
DB 874 cggcccggtacagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 933
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 733 CGGCGCGGCAATTCATTTGACACTTCGCTCGAGAGGCGGCAATATATCTGT 792
DB 934 aaccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 993
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 793 TTGACCGGATGGAGCGGCTACTGTACATACCAATCGCCCAAGGCTGGA--A- 849
DB 994 cagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 1053
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 850 CATTG--GACAT--TCCGCAACCAATCAATGCTTTGCACTGAGGCTGGAAGGAGCC 906
DB 1054 tactgtgagcagagtgatccatccagcgagcgagcgagcgagcgagcgagcgagcgag 1113
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 907 TTGTGACACAGTGGCTTTCGAAACCGAAGTGCATGACGACCTCAACAAACCTCTGG 966
DB 1114 cagtgcaacgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 1173
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 967 ACTGTGATCTGCAACGAGGCTTGGGAGGCTTGTACTGCAACAGGATCTGAATCTATGCTC 1026
DB 1174 actccacacagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 1199
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 1027 ACCAACACAGACCCCTGCAAGATGG 1052

RESULT 5
ID T70174 standard; cDNA to mRNA; 2663 BP.
AC T70174.
DT 11-FEB-1998 (first entry)
DE Proliferation and differentiation suppression protein encoding cDNA.
KW Proliferation; differentiation; suppression; human; delta-1;
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression; ds.
OS Homo sapiens.
FH Key location/Qualifiers
FT CDS 179..2350
FT FT /*tag= a
FT FT /product= Differentiation-suppression_protein
FT FT sig_peptide 179..241
FT FT /*tag= b
FT FT mat_peptide 242..2347
FT FT /*tag= c
FN W09719172-A1.
PD 29-MAY-1997.
PE 15-NOV-1996; J03356.
PR 30-NOV-1995; JP-311811.
PR 17-NOV-1995; JP-299611.
PA (ASAH) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR P-PSDB; M18353.
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 15; Page 77-82; 114pp; Japanese.
CC The present sequence encodes a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 2663 BP; 534 A; 825 C; 829 G; 475 T;

Query Match 4.5%; Score 131; DB 34; Length 2663;

Best Local Similarity 61.8%; Pred. No. 1.43e-66;
Matches 514; Conservative 0; Mismatches 305; Indels 13; Gaps 12;

DB 318 gggcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 376
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 295 GGACGAGGCGCAAGTGGCTGGGAGCTGCAAGACGCGGTTCCGCTGCTTAAGCAC 354
DB 377 taccaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 436
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 355 TACAGGCGCAATTCAGACCACTCCAGTGCACCTACCTGAGGAGATGATACACGCAAT 414
DB 437 ctggcggtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 493
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 415 CTGCGCGGAAGTGGCTCAATGACCAACGCGCCAGCCCTTCAGAAAGGCTTCACG 474
DB 494 aaccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 553
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 475 AATCCATTCAGTCCCTTTCGTTTCATGAGCGGAGTACCTTTCCTATGCTCGAG 534
DB 554 gctctcacacagattctcctgagctcgcagacagacagacagacagacagacagacagac 613
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 535 GC-CTGG-CAT-GATACCAACATAGCGGCAATGCGCAACCAACCTCTATCCAG 591
DB 614 cgcctgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 673
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 592 CGACTCTGTGGTACGAGGCTGAGGCTGCTCCGAAATGAGAGCAACAAAGCGGAA 651
DB 674 agcgcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 733
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 652 TCGAGTACAGTCCGCTGAGTACGATTTCCGTACCTGCGATTCACACTACTACGGA 711

Db 734 gagggctgctcgtttctgcccgcgggacgagtccttgcggccacttccactcgtggtg 793
QY 712 TCCGCTGTGCCAAGTTCGCGCGCCCGCCGACGANTCATTTGGACACTCGACTGCTCG 771
Db 794 gaggctgggagaaagtgtgaacacctggtctggaaggccctacttgcacagagcgatc 853
QY 772 GAGCGGGGAAATATCTGTTTGACCGGATGCGAGGCGATTACTGTACATACCANA 831
Db 854 tgcctgctggtgtgtagagcagcatgatttggtaaaacccaggggaatgaagtgc 913
QY 832 TGC--GCCAA-A-G-GCTGTGA-CATGGACATTGGACAAACCAATCAATCGTTTC 885
Db 914 aagtggtgctgagggcggtactgtgacagtgatcgcgtatccaggtctccat 973
QY 886 CAATCGGGTGAAGGAGCCCTGTGCAACGAGTGGCTTCTGGAACCGAACTGCATCCAT 945
Db 974 ggcactgcccagcagcggcagtcgagtcgcaactgcccaggaaggctgggggcccctttctgc 1033
QY 946 GGCACCTGCAACAAACCTGGACTTGCATCTGCAACGAGGTTGGGAGGCTTGTACTGC 1005
Db 1034 aaccggacctgaactgacacacacataagccctgcaagaatgagccacctgacc 1093
QY 1006 AACCGAGGATCTGAATCTGCAACCAACAGACACCTGCAAGACCTGCGGAACCTGCTC 1065
Db 1094 aacacggcgagggagctacacttctcttgcggccctgggtacacaggtg 1145
QY 1066 AACACGGCGAGGATTGTACATGCAATGCAATGCTCCAGGATACAGTGTTG 1117

RESULT

ID T58899 standard; cDNA; 2692 BP.
AC T58899;
DT 28-APR-1997 (first entry)
DE M-Delta-1 gene.
KW M-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;
KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy; ss.
OS Mus sp.
FH Key Location/Qualifiers
FT cds 34..2202 /*tag= a
PN WO9701571-A1.
PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI (UYTA) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
PI Lewis J;
DR WPI; 97-100159/09.
DR P-PSDB; W11720.
PT New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PS Disclosure; Fig 7A-7B; 135pp; English.
CC The M-delta-1 gene (T58899) codes for the mouse homologue (W11720)
CC of Drosophila Delta, a protein that binds to Notch protein. It was
CC obtd. by PCR amplification of cDNA derived from 8.5 to 9.5-day
CC mouse embryos using primers (see also T59457-58) based on Delta
CC and EGF-like repeat sequences. M-Delta-1 is primarily expressed
CC in the presomitic mesoderm, central nervous system, peripheral
CC nervous system, and kidney. Chick (T58897-98) and human (T58900,
CC T59454) Delta-1 sequences have also been isolated. Delta-1 genes
CC can be used in the prodn. of Delta polypeptides and (including
CC antisense sequences) utilised in the treatment of disorders of cell
CC fate or differentiation, such as cancer, and nervous system
CC disorders, or to promote tissue regeneration and repair.
SQ Sequence 2692 BP; 592 A; 748 C; 768 G; 581 T;
Query Match 3.7%; Score 108; DB 28; Length 2692;
Best Local Similarity 61.6%; Pred. No. 9,37e-53;
Matches 498; Conservative 0; Mismatches 294; Indels 16; Gaps 13;

Db 195 ctgcagagaccttcttccgctatgctcctcaagcactaccagccagcggtgtcaccggagcc 254
QY 321 CTGCAAGACGGGTTCGCGTCTGCTTAAGACACTACCAGGCCACCATCGACACCACTTC 380
Db 255 acctgcactacgagcgagtcgcgcagtcgcgcagtcggtgtgctgactcctcagcctgc 314
QY 381 CAGTGCACCTAGGGGACGTGATCAGGCCATCTTCGGCGAGAACTCGGTCAATCTGAC 440
Db 315 tgatgctgcagcgatc--gacccgc-cttcagcaacccctccatccgattccccctggct 371
QY 441 CGAGCCCGGCGCTTCAGAAACAAGGCTTCAGAAATCCATCCAGTTCCTCCCTCTCGTT 500
Db 372 caactggcaggtacattctctctgataattgaagccctccatacagactctcccgatga 431
QY 501 CTCATGGCGGGTACCTTCTCGCTGATGCTGAGGCTTGGCATG-ATAGAAACAATAGCG 559
Db 432 cctcgcacaagaaacccagaaagactcatcagccctcgtgacccacacagagggcacctcac 491
QY 560 GCAATCGGC- GAACCAACA- AGCTCTCTATCCAGCGACTCTTGGTCAGCAGTACTGGA 617
Db 492 tgtgggagaagaatggtctcagggact- tcaagtagcgcgcgcacacagacctccggtact 550
QY 618 GGTGCTCTCGAATGGAAGACGAACAAGTCGGAATCGCAGTA-CACGTCTCGTGGAGTACG 676
Db 551 cttaacgggttgtgtgacgagcactactacgagaaggtgtctgtgttctgcgcac 610
QY 677 ATTTCTGCTGTCACCTGCGATCTCAACTACTACGATCCGGCTGTGCCAAGTTCTGCGGC 736
Db 611 ctgggagatcgccttggccacttccactgcggggacagaggggagagatgctgcacc 670
QY 737 CCGCGACGATTCATTTGGACACTCGACTGTCTCGAGACGGCGGAAATATCTGTTGA 796
Db 671 ctggctggaagccagtagtctg- cactgaacccaatgtctgtccaggggtgtgatcccaa 729
QY 797 CCGGATGGCAGGCGATTACTGTACATACCCAAA-TG-C-GCCAAAGGCTG-TGA--A 849
Db 730 catgatactgtgacaaacccagggagtgcaagtgcagagttgctgctggcagggccctac 789
QY 850 CATGGACATGGCACAACCAACCAATCAATCGTTTGCACACTGGGCTGGGAAGGAGCTTG 909
Db 790 tgcgagtagtgcacccgataccaggtgtgtctccatggccactggccagcaacctggcag 849
QY 910 TGCACGAGTGGCTTCTGGNACCGAACTGCATCTCATGGCACCCTGCAACAACCTCGACT 969
Db 850 tgtaactgcaggaagctggtggggccctttctgcaacaaagacctgactactgtact 909
QY 970 TGCATCTGCAACGAGGTTGGGGAGGCTTGTACTGCAACCAAGGATCTGAACCTACTGCACC 1029
Db 910 caccataagcgtgaggaatgagccactgcaacacacacagggccagggaggaactacaca 969
QY 1030 AACCAAGACCCCTGCAAGAAATGGCGAATGCTTCAACACCGCGGAGGATGTATACACA 1089
Db 970 tgtctctgcgcagcctgggtatatacaggtg 997
QY 1090 TGCAAATGGCTCCAGGATACAGTGGTG 1117

RESULT

ID T59454 standard; DNA; 1980 BP.
AC T59454;
DT 28-APR-1997 (first entry)
DE H-Delta-1 contig clone Hdl18.
KW H-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;
KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy; ss.
OS Homo sapiens.
PN WO9701571-A1.
PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.


```
Best Local Similarity 60.1%; Pred. No. 1.23e-16;
Matches 247; Conservative 0; Mismatches 158; Indels 6; Gaps 4;

Db 946 gagtaccagatccggtgacctgtagtactactactactggttggcgtgcaataagttc 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 670 GAGTACGATTTCCTCGTCACTCGCATCTCACTACTAGGATCGGCTGTGCCAAGTTC 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1006 taccgccccagagatgactcttttgacactatgctgtgaccagaatggcaacaaact 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 730 TGCCGGCCCCGGACGATTCAITTTGGACACTCGACTTGTCTGGAGACGGCGGAATATC 789
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1066 tcatggaaggctggatggccccgaatgtaacagagctatttgcgcagaagctgcagt 1125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 790 TGTTTGACCGGATGGCAGGCGATTTACTGTCATATACCCAAATGC-GCCAAAG--GCTGT 846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1126 cctagcagctgtcttgcgaactcccgagtgcagtgcaagtgccagtcagcgtggcaagc 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 847 G--AA-CATGGACATTTCGACAAACCAATCAATGCGTTTGCCAACTGGGCTGGAAGGGA 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1186 ctgtactgtataagtgcacacacccggagatgcgtccagggcatctgtaatgagccc 1245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 904 GCGTTGTGCAAGAGTGCCTCTGGAACCGCACTGCTCATCCATGGCACCTGCAACAACCC 963
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1246 tggcagtcctctgtgagaccactggcgccagctctgtgacaaagatctcaattac 1305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 964 TGGACTTGCATCTGCAACGAGGTTGGGGAGGCTTGTACTGCAACGAGGATCTGAACTAC 1023
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1306 tgtggagactcagccgtgtctcaacggcggaactgttagcaacacagc 1356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1024 TGCACCAACACAGACCCCTGCAAGATGCGGAACCTGCTTCAACACCGGC 1074
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
ID T40091 standard; cDNA; 4483 BP.
AC T40091;
DE 28-JAN-1997 (first entry)
DT Human Serrate-2 (HJ2) cDNA (containing an internal deletion).
KW Serrate-2; human jagged-2; HJ2; Notch; cell differentiation;
KW cell fate; central nervous system; cancer; tissue repair; therapy;
KW diagnosis; ds.
OS Homo sapiens.
PI Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
FH Key 332.4105
FT cds
FT misc_difference 240..241
FT /*tag= a
FT /*tag= b
FT /*note= "there is a deletion of approx. 120
FT nucleotides in this region of the sequence,
FT probably resulting from a cloning artifact
FT in construction of the library, encoding a
FT portion of the signal peptide and beginning
FT of the DSL domain"
PN W09627610-A1.
PD 12-SEP-1996.
PE 07-MAR-1996; U03172.
PF 07-MAR-1995; US-400159.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYVA ) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
PI Lewis JH, Mann RS, Myat AM;
DR WPI: 96-425379/42.
DR P-PSDB; W05834.
PT Vertebrate Serrate protein and related DNA - used to treat or
PT prevent malignancies characterised by increased Notch activity.
PS Disclosure: Page 98-103; 161pp; English.
CC 2 cDNA clones (T40090 and T40091) respectively code for human
CC Serrate-1 (W05833) and human Serrate-2 (W05833), ligands for Notch
CC that are believed to play a major role in determining cell fates in
CC the central nervous system. The clones were isolated from a foetal
CC brain cDNA library using probes obtd. by PCR amplification of human
CC placental cDNA with primers based on conserved sequences of
CC Drosophila Serrate and Delta. The cDNA can be incorporated into
CC vectors and utilised in the treatment and diagnosis of disorders of
```

```
CC cell fate and differentiation; antisense sequences can be used to
CC treat malignancies of cell types that express Serrate or Notch.
SQ Sequence 4483 BP; 780 A; 1389 C; 1472 G; 842 T;

Query Match 1.8%; Score 51; DB 24; Length 4483;
Best Local Similarity 64.1%; Pred. No. 2.09e-15;
Matches 116; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Db 620 ggcgtgcaaggagggttctgcgatgagtggtccctaccctaccgggtgctgcatggcagt 679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 892 GCGTGAAGGGAGCCCTGTGCAACGATGCGTTCTGGAACCGAAGTGCATCTCCATGGCACC 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 680 tgggtgagccctggcagtgcaactgtgacacaaactggggcgccctgctctgtgacaaa 739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 952 TGCACAACCCCTGGACTTGCATCTGCACGAGGGTGGGAGGCTTGTACTGCAACCCAG 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 740 gacctgaactactgtggcagccaccacccctgcacaaacggaggcagctgcatcaacgcc 799
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1012 GATCTGAATCTACTGCAACCAACACAGACCCCTGCAAGAATGGCGGAACCTGCTTCAACACC 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 800 g 800
Qy 1072 G 1072

RESULT 11
ID Q43910 standard; DNA; 5561 BP.
AC Q43910;
DE 30-NOV-1993 (first entry)
DT Sequence encoding a serrate protein.
KW Serrate; topoythmic protein; family; ss.
OS Drosophila melanogaster.
FH Key Location/Qualifiers
FT cds 442..4653
FT /*tag= a
PN W09312141-A.
PD 24-JUN-1993.
PF 11-DEC-1991; U09240.
PR 11-DEC-1991; WO-009240.
PA (UYVA ) UNIV YALE.
PI Artavanis-Tsakonas S, Fleming RJ;
DR WPI: 93-214095/26.
DR P-PSDB; R38304.
PT Purified serrate protein, nucleic acid and antibodies - used in
PT the study and manipulation of differentiation and other
PT physiological processes
PS Claim 39; Pages 74-80; 119pp; English.
CC Two Drosophila genomic phage libraries were screened and recombinant
CC clones were isolated. The cDNAs in lambda gt10 were isolated from an
CC early pupal library. The C1 cDNA was isolated from an early pupal
CC library. Subsequently the C3 cDNA was isolated using the 5' 700 bp
CC terminal fragment of the C1 cDNA as probe. The complete 5561bp
CC sequence of DNA of the Drosophila Serrate protein was derived from
CC C1 and C3 cDNAs (Q43910). The deduced protein product appears to be
CC a transmembrane protein. AAs 51-80 represent the likely signal
CC peptide; aAs 542-564 represent the putative transmembrane domain.
CC aAs 1221-1245 represent the putative transmembrane domain.
SQ Sequence 5561 BP; 1398 A; 1447 C; 1504 G; 1212 T;

Query Match 1.8%; Score 52; DB 7; Length 5561;
Best Local Similarity 66.7%; Pred. No. 5.09e-16;
Matches 134; Conservative 0; Mismatches 64; Indels 3; Gaps 3;

Db 1340 aatgcgaatgcagaccggcgtgctggttcattgtgcaacgagtgcatggtctatcccg 1399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 875 AATGCGTTTGCACACTGGGCTGGAAGGAGCCCTGTGCAACGAGTGGTCTTGTGGAACCGA 934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1400 gctgcgaagcagtggttctcccaacggcagcgcctggaataatgctgtgcaacacaaatggg 1459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 935 ACTGCTCATCCATGGCACCTGCAACAA-A-C-CTTGGACTTGCATCTGTCACAGGGTGGG 991
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1460 gtggcatattgtgcataagatttaattctcgccaccctgaacctgaacacgacg 1519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

M P S E R I H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jun 13 10:07:22 1998: Maspar time 126.49 Seconds

Tabular output not generated. 1205.088 Million cell updates/sec

Title: >US-08-083-590A-1
Description: (1-2892) from US08083590A.seq
Perfect Score: 2892
N.A. Sequence: 1 GAATCGAGGAGATATATCA.....CGCGTTTCGATCCGGAATC 2892
Comp: CTTAGGCTCTTATATAGT.....CGCGAAGCTAGGGCTTAG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 102132 seqs, 26354270 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-Issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 9.454; Variance 5.035; scale 1.878

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 2892 | 100.0 | 2892 | 1 | US-08-264- Sequence 5, Applicatio | 0.00e+00 |
| 2 | 71 | 2.5 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 5.66e-32 |
| 3 | 54 | 1.9 | 8378 | 2 | PCT-US91-0 Sequence 1, Applicati | 5.18e-20 |
| 4 | 42 | 1.5 | 215 | 1 | US-08-238- Sequence 5, Applicatio | 4.96e-12 |
| 5 | 42 | 1.5 | 1573 | 1 | US-08-597- Sequence 4, Applicatio | 4.96e-12 |
| 6 | 42 | 1.5 | 1573 | 1 | US-08-457- Sequence 4, Applicatio | 9.50e-11 |
| 7 | 40 | 1.4 | 215 | 1 | US-08-238- Sequence 3, Applicatio | 7.35e-09 |
| 8 | 37 | 1.3 | 1556 | 1 | US-08-457- Sequence 3, Applicatio | 7.35e-09 |
| 9 | 37 | 1.3 | 1556 | 1 | US-08-597- Sequence 3, Applicatio | 7.35e-09 |
| 10 | 38 | 1.3 | 7218 | 1 | US-08-232- Sequence 14, Applicati | 1.74e-09 |
| 11 | 32 | 1.1 | 4255 | 2 | PCT-US96-0 Sequence 9, Applicatio | 8.10e-06 |
| 12 | 32 | 1.1 | 4835 | 2 | PCT-US96-0 Sequence 8, Applicatio | 1.67e-03 |
| 13 | 28 | 1.0 | 1320 | 1 | US-08-264- Sequence 1, Applicatio | 1.67e-03 |
| 14 | 28 | 1.0 | 1425 | 1 | US-08-464- Sequence 1, Applicatio | 1.67e-03 |
| 15 | 28 | 1.0 | 1425 | 1 | US-08-385- Sequence 1, Applicatio | 1.20e-04 |
| 16 | 30 | 1.0 | 5191 | 1 | PCT-US93-0 Sequence 1, Applicatio | 1.20e-04 |
| 17 | 30 | 1.0 | 5191 | 1 | US-08-485- Sequence 100, Applicat | 7.35e-02 |
| 18 | 25 | 0.9 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 7.35e-02 |
| 19 | 25 | 0.9 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 7.35e-02 |
| 20 | 25 | 0.9 | 74 | 2 | PCT-US95-1 Sequence 94, Applicati | 7.35e-02 |

| | | | | | | | |
|----|----|-----|-------|---|------------|-------------------------|----------|
| 21 | 25 | 0.9 | 81 | 2 | PCT-US95-1 | Sequence 92, Applicati | 7.36e-02 |
| 22 | 25 | 0.9 | 81 | 2 | PCT-US95-1 | Sequence 98, Applicati | 7.36e-02 |
| 23 | 26 | 0.9 | 81 | 2 | PCT-US95-1 | Sequence 97, Applicati | 7.36e-02 |
| 24 | 25 | 0.9 | 82 | 2 | PCT-US95-1 | Sequence 13, Applicatio | 6.02e-03 |
| 25 | 26 | 0.9 | 105 | 1 | US-07-865- | Sequence 1, Applicatio | 7.36e-02 |
| 26 | 27 | 0.9 | 242 | 1 | US-08-273- | Sequence 163, Applicat | 7.36e-02 |
| 27 | 25 | 0.9 | 242 | 1 | US-08-273- | Sequence 163, Applicat | 7.36e-02 |
| 28 | 25 | 0.9 | 745 | 2 | PCT-US95-0 | Sequence 163, Applicat | 7.36e-02 |
| 29 | 25 | 0.9 | 745 | 1 | US-08-036- | Sequence 163, Applicat | 7.36e-02 |
| 30 | 25 | 0.9 | 745 | 1 | US-08-249- | Sequence 163, Applicat | 7.36e-02 |
| 31 | 26 | 0.9 | 1201 | 3 | 5252556-2 | Patent No. 5252556. | 2.13e-02 |
| 32 | 25 | 0.9 | 1383 | 1 | US-08-484- | Sequence 1, Applicatio | 7.36e-02 |
| 33 | 25 | 0.9 | 2003 | 2 | PCT-US94-0 | Sequence 21, Applicati | 7.36e-02 |
| 34 | 25 | 0.9 | 2003 | 1 | US-08-036- | Sequence 21, Applicati | 7.36e-02 |
| 35 | 25 | 0.9 | 2003 | 1 | US-08-249- | Sequence 21, Applicati | 7.36e-02 |
| 36 | 25 | 0.9 | 2003 | 1 | US-08-469- | Sequence 21, Applicati | 7.36e-02 |
| 37 | 25 | 0.9 | 2950 | 2 | PCT-US93-0 | Sequence 1, Applicatio | 7.36e-02 |
| 38 | 25 | 0.9 | 2968 | 2 | PCT-US93-0 | Sequence 1, Applicatio | 7.36e-02 |
| 39 | 25 | 0.9 | 4978 | 1 | US-08-220- | Sequence 1, Applicatio | 7.36e-02 |
| 40 | 26 | 0.9 | 5288 | 2 | PCT-US95-1 | Sequence 18, Applicati | 7.36e-02 |
| 41 | 27 | 0.9 | 5775 | 1 | US-08-306- | Sequence 15, Applicati | 6.02e-03 |
| 42 | 27 | 0.9 | 5775 | 2 | PCT-US93-0 | Sequence 25, Applicati | 6.02e-03 |
| 43 | 25 | 0.9 | 11613 | 1 | US-08-484- | Sequence 10, Applicati | 7.36e-02 |
| 44 | 26 | 0.9 | 35100 | 2 | PCT-US93-0 | Sequence 19, Applicati | 2.13e-02 |
| 45 | 26 | 0.9 | 35100 | 1 | US-08-306- | Sequence 19, Applicati | 2.13e-02 |

ALIGNMENTS

RESULT 1
ID US-08-264-534-5 STANDARD; DNA; UNC; 2892 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5, Application US/08264534.
CC Sequence 5, Application US/08264534
CC Patent No. 5648464
CC GENERAL INFORMATION:
CC APPLICANT: Attavans-Teakonas, Spyridon et al.
CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/264,534
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/695,189
CC FILING DATE: 03-MAY-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 7326-004
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2892 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double

| | | | |
|----------|-----------------------------------------------|----------------------------------------------------------------|------|
| Db | 2041 | GGAGGAACACAGAACCAACGATGGCGGTGGCCCAATGCATGCATATGGCAGTGGGGT | 2100 |
| Qy | 2041 | GGCAGGAACACAGAACCAACGATGGCGGTGGCCCAATGCATGCATATGGCAGTGGGGT | 2100 |
| Db | 2101 | GGGTGTAGCTTTGGCTTCAGGCTCTCTGGGGGGCAAACTGGCAGCAACAGCGGTCTACC | 2160 |
| Qy | 2101 | GGGTGTAGCTTTGGCTTCAGGCTCTCTGGGGGGCAAACTGGCAGCAACAGCGGTCTACC | 2160 |
| Db | 2161 | TTTCGATGGCGGCAACCCGAATATCATCAAAACACCTGGGAGCAAGTGGTCAACACATT | 2220 |
| Qy | 2161 | TTTCGATGGCGGCAACCCGAATATCATCAAAACACCTGGGAGCAAGTGGTCAACACATT | 2220 |
| Db | 2221 | TTTGCCCTCAGACAGCAGCGGGGGGGGGGCGACAGCGGGCGACAGTGTCTCATATAC | 2280 |
| Qy | 2221 | TTTGCCCTCAGACAGCAGCGGGGGGGGGGCGACAGCGGGCGACAGTGTCTCATATAC | 2280 |
| Db | 2281 | GGCGGATATGTGGCTCTCGGTGGGGATTAACAACAATGCCAACTCAGACTTTTGTGGCT | 2340 |
| Qy | 2281 | GGCGGATATGTGGCTCTCGGTGGGGATTAACAACAATGCCAACTCAGACTTTTGTGGCT | 2340 |
| Db | 2341 | CCGCTTACAAAGAGCCAAAGTCGCAAAAAGCACTAACACCGATCCACGCTCATGACCCG | 2400 |
| Qy | 2341 | CCGCTTACAAAGAGCCAAAGTCGCAAAAAGCACTAACACCGATCCACGCTCATGACCCG | 2400 |
| Db | 2401 | GGTTGCCCGGACAGGAGCTCAGCAAGGGAGCGCTGCGGAGAGACCGGGAGCGGCGGAG | 2460 |
| Qy | 2401 | GGTTGCCCGGACAGGAGCTCAGCAAGGGAGCGCTGCGGAGAGACCGGGAGCGGCGGAG | 2460 |
| Db | 2461 | GGCAAGAGGATCTGTGTTTGGCGAGGGTCTCTACTGTAGCCAGCGTTGGCCCTCGTG | 2520 |
| Qy | 2461 | GGCAAGAGGATCTGTGTTTGGCGAGGGTCTCTACTGTAGCCAGCGTTGGCCCTCGTG | 2520 |
| Db | 2521 | GGGGGGGGGGAGTGGCCCGGAGGCTTTCAATCCAGCTAATGGCTGCAGCTTGGCGACG | 2580 |
| Qy | 2521 | GGGGGGGGGGAGTGGCCCGGAGGCTTTCAATCCAGCTAATGGCTGCAGCTTGGCGACG | 2580 |
| Db | 2581 | GGCAGCGGAGCGGGGAGCGGCGCAACAGCAGATCGGTGCTGCGGCACTCCGCAATATG | 2640 |
| Qy | 2581 | GGCAGCGGAGCGGGGAGCGGCGCAACAGCAGATCGGTGCTGCGGCACTCCGCAATATG | 2640 |
| Db | 2641 | TAACTCCAAAAATCCGGAAAGGCTCCTGGTAAATCCGGAGAAATCCGCATGAGAGACTG | 2700 |
| Qy | 2641 | TAACTCCAAAAATCCGGAAAGGCTCCTGGTAAATCCGGAGAAATCCGCATGAGAGACTG | 2700 |
| Db | 2701 | AACAGCATTACCAAGAAAGAAAGCTGGGTGGTTCAAAATGTGAGAGAAACGCAAAAT | 2760 |
| Qy | 2701 | AACAGCATTACCAAGAAAGAAAGCTGGGTGGTTCAAAATGTGAGAGAAACGCAAAAT | 2760 |
| Db | 2761 | GTTGTTGTTGATTGAAGCAGTTAGTGCATCCAAAAATGAAAAATCTGTAAAGAGCATA | 2820 |
| Qy | 2761 | GTTGTTGTTGATTGAAGCAGTTAGTGCATCCAAAAATGAAAAATCTGTAAAGAGCATA | 2820 |
| Db | 2821 | ACTCGTAAACTCCCTTAAAAAATTTGTATAGTAATTAAGCAAAAGCTGTGACCCAGCGCTTC | 2880 |
| Qy | 2821 | ACTCGTAAACTCCCTTAAAAAATTTGTATAGTAATTAAGCAAAAGCTGTGACCCAGCGCTTC | 2880 |
| Db | 2881 | GATCCCGAATTC 2892 | |
| Qy | 2881 | GATCCCGAATTC 2892 | |
| RESULT 2 | | | |
| ID | US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP. | | |
| AC | XXXXXX | | |
| DT | 01-JAN-1900 | | |
| DE | Sequence 14, Application US/08332463. | | |
| CC | Sequence 14, Application US/08332463 | | |
| CC | Patent No. 5670367 | | |
| CC | GENERAL INFORMATION: | | |
| CC | APPLICANT: DORNER, F. | | |
| CC | APPLICANT: SCHEIFLINGER, F. | | |
| CC | APPLICANT: FALKNER, F. G. | | |

```
CC TITLE OF INVENTION: RECOMBINANT FOVLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEX: (703)683-4109
CC FAX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZgpt-F15
CC SEQ Sequence 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;

Query Match      2.5%; Score 71; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 6.6e-32;
Matches    12; Conservative   215; Mismatches 156; Indels     0; Gaps     0;
```


Best Local Similarity 14.8%; Pred. No. 4.96e-12;
Matches 31; Conservative 87; Mismatches 88; Indels 3; Gaps 2;

Db 6 MSSSVSRASCNDRKAKKNGNTSSMTTDOCNFTMGVCDITTYRVNNDGHNKISSAN 65
Cc 498 CGAAGAGGGAAGTCTGATGGATTCGGAAGCCCTGTCTTGAAGAGCGGTGGCGT 439
Db 66 YN-C-YGGNNVGAATHYHTNVSGADSKTYTDS-YNAGTSSNGGTGDRSGADSYGS 122
Cc 438 CAATATACCGAGTCTTCGCCGAGATGCGCGTATCAGTCCCGTAGTGACACTGGGA 379
Db 123 SKTAMTSRNTGKTANNAYDSVRNMGDAVSQDKNTKKHAKNSADQKYSKNNGDRNNRYG 182
Cc 378 GGTGTGTCTGATGCTGCGCTGTAGTGTCTTAGCGACAGCGGAAACCGGCTTGACAGCT 319
Cc 183 TGTSMVSNNGCGGNKADSVSYANNKCCG 211
Cc 318 GCCAGGCACTTGCCCTCGCTCGTCCG 290

RESULT 5
ID US-08-597-545-4 STANDARD; DNA; UNC; 1573 BP.

AC xxxxxx
DE 01-JAN-1900
Sequence 4, Application US/08597545.
Cc Sequence 4, Application US/08597545
Cc Patent No. 5580738
Cc GENERAL INFORMATION:
Cc APPLICANT: LABORDA, Jorge
Cc TITLE OF INVENTION: Delta-Like Gene Expressed In
Cc TITLE OF INVENTION: Neuroendocrine Tumors
Cc NUMBER OF SEQUENCES: 9
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Foley & Lardner
Cc STREET: 3000 K Street, N.W., Suite 500
Cc CITY: Washington, D.C.
Cc COUNTRY: USA
Cc ZIP: 20007-5109
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patentin Release #1.0, Version #1.25
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: US/08/597,545
Cc FILING DATE:
Cc CLASSIFICATION: 435
Cc PRIOR APPLICATION DATA:
Cc APPLICATION NUMBER: US 07/989,537
Cc FILING DATE: 11-DEC-1992
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: BENT, Stephen A.
Cc REGISTRATION NUMBER: 29,768
Cc REFERENCE/DOCKET NUMBER: 40399/166 NIHD
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (202)672-5300
Cc TELEFAX: (202)672-5399
Cc TELEX: 904136
Cc INFORMATION FOR SEQ ID NO: 4:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 1573 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: double
Cc TOPOLOGY: linear
Cc ORIGINAL SOURCE:
Cc STRAIN: Mouse Dlx
Cc Sequence 1573 BP; 327 A; 507 C; 419 G; 320 T; 0 other;

Query Match 1.5%; Score 42; DB 1; Length 1573;
Best Local Similarity 58.3%; Pred. No. 4.96e-12;
Matches 200; Conservative 0; Mismatches 140; Indels 3; Gaps 2;

Db 234 ATGATTTGCGAGGCTGACATGTCTGACAGTGCCATGTGCTGGAGGCTCCCTCT 293

Qy 851 ATGACATTTGCGAAGAAACCAATCATGCTTTGGCACTGGGCTGGAAGGAGCTTGT 910
Db 294 GTGACAGATGTGTACTACCCCTGGCTGTGCATATGAGTGTGCAAGAACCATGGAGCT 353
Qy 911 GCAACGAGATGGTCTGTGAAACGCAATGCATCATCGGACCTGCAAAACCCCTGACTT 970
Db 354 GCATCTGCAAGATAGCTGGGAGCGGAAATCTGCGAATATGACGTTGGGCTTGACACT 413
Qy 971 GCATCTGCAAGAGGTTGGGAGGCTGTACTGCAACCAAGATCTGACTACTACACCA 1030
Db 414 -CA-ACCCCTGGCCCAACAATGAACTGCGTGGACCTGGAGAAAGCCAGTACGAAT 470
Qy 1031 ACCACAGACCTTGAAGAATGCGGAACCTGCTTCAACACCGGAGGAGATTGTACACAT 1090
Db 471 GCTCTGCAACACTGGGTTCTCTGGAAGACTGCCACACAGGCTGGCCCTGCGTGA 530
Qy 1091 GCAATGCGCTTCAGAGATCAATGATGATGATGATGATGATGATGATGATGATG 1150
Db 531 TCAATGATTCCTCGCCAGCAGGAGGCGCTGGGTGATGA 573
Qy 1151 CCGATGTCAATCCTCGCCAGATGATGATGATGATGATGATGATGATGATGATG 1193

RESULT 6
ID US-08-457-135-4 STANDARD; DNA; UNC; 1573 BP.

AC xxxxxx
DE 01-JAN-1900
Sequence 4, Application US/08457135.
Cc Sequence 4, Application US/08457135
Cc Patent No. 5644031
Cc GENERAL INFORMATION:
Cc APPLICANT: LABORDA, Jorge
Cc TITLE OF INVENTION: Delta-Like Gene Expressed In
Cc TITLE OF INVENTION: Neuroendocrine Tumors
Cc NUMBER OF SEQUENCES: 9
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Foley & Lardner
Cc STREET: 3000 K Street, N.W., Suite 500
Cc CITY: Washington, D.C.
Cc COUNTRY: USA
Cc ZIP: 20007-5109
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patentin Release #1.0, Version #1.25
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: US/08/457,135
Cc FILING DATE: 01-JUN-1995
Cc PRIOR APPLICATION DATA:
Cc APPLICATION NUMBER: 07/989,537
Cc FILING DATE: 11-DEC-1992
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: BENT, Stephen A.
Cc REGISTRATION NUMBER: 29,768
Cc REFERENCE/DOCKET NUMBER: 40399/304/NIHD
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (202)672-5300
Cc TELEFAX: (202)672-5399
Cc TELEX: 904136
Cc INFORMATION FOR SEQ ID NO: 4:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 1573 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: double
Cc TOPOLOGY: linear
Cc ORIGINAL SOURCE:
Cc STRAIN: Mouse Dlx
Cc Sequence 1573 BP; 327 A; 507 C; 419 G; 320 T; 0 other;

Query Match 1.5%; Score 42; DB 1; Length 1573;
Best Local Similarity 58.3%; Pred. No. 4.96e-12;

| | | | |
|--------|--------------------------------------------------------|-------------------------------------------------------------------|------|
| Db | 271 | AAATGGATTGTGGAGAGATGCAAAATGTTTGACAGGTGCCATGTGCGGTGGCAGGGGTCCC | 330 |
| Db | 848 | AAATGGATTGTGGAGAGATGCAAAATGTTTGACAGGTGCCATGTGCGGTGGCAGGGGTCCC | 330 |
| Qy | 848 | AAATGGATTGTGGAGAGATGCAAAATGTTTGACAGGTGCCATGTGCGGTGGCAGGGGTCCC | 330 |
| Db | 331 | TTTGTGACCAATGTGGTGTGACCTCTCTCCGGCTGACCTTACAGGACTCTGTGGAGAACCCTGGC | 390 |
| Qy | 908 | TGTGCAACGAGATGTGGCTTCTGGAACCGAAGCTGATTCATGAGCACTGTGCAACCAACCTGGA | 967 |
| Db | 391 | AGTGCAATTTGCACGACGCGCTGGGACGGGAGCTCTGTGTATAGATGTTTGGGCTGCT | 450 |
| Qy | 968 | CTTGCAATCTGCAAGAGAGGTTGGGAGGCTTGTACTGTGCAACAGAGATCTGAACTACTGCA | 1027 |
| Db | 451 | CCT-CCGGC-CCCTGTGGCAACAAGGGGACCTGCGTAGGCTGGAGGCTGGCTCTATTG | 507 |
| Qy | 1028 | CCAAACACAGACCTGTGCAAGAAATGGCGGAACCTGTCTTCAACACCGGCGAGGATTGTACA | 1087 |
| Db | 508 | AATGCTCTCTGTGCCCCCGGGTACTCGGGAAGACCTGCCAGAAAAGCAGCGGCCCTGTG | 567 |
| Qy | 1088 | CATGCAATGTGGCTTCCAGAGATACAGTGTGTATTTTGTCCAAATGAGATCTACTCTGG | 1147 |
| Db | 568 | TGATCAACGGCTCCCTCTGCCACGACGAGGAGCACTGCGTGATGA | 613 |
| Qy | 1148 | ATGCGAGTGTCAATCCTGCGCAGAAATGGTGTACTGTGCAACGAGATCTGCAATGATGA | 1193 |
| RESULT | 9 | | |
| ID | US-08-597-545-3 | STANDARD; DNM; UNC; 1556 BP. | |
| AC | xxxxxx | | |
| DT | 01-JAN-1900 | | |
| CC | Sequence 3, Application US/08597545. | | |
| CC | Sequence 3, Application US/08597545 | | |
| CC | Patent No. 5580738 | | |
| CC | GENERAL INFORMATION: | | |
| CC | APPLICANT: LABORDA, Jorge | | |
| CC | TITLE OF INVENTION: Delta-like Gene Expressed In | | |
| CC | TITLE OF INVENTION: Neuroendocrine Tumors | | |
| CC | NUMBER OF SEQUENCES: 9 | | |
| CC | CORRESPONDENCE ADDRESS: | | |
| CC | ADDRESSEE: Polley & Lardner | | |
| CC | STREET: 3000 K Street, N.W., Suite 500 | | |
| CC | CITY: Washington, D.C. | | |
| CC | COUNTRY: USA | | |
| CC | ZIP: 20007-5109 | | |
| CC | COMPUTER READABLE FORM: | | |
| CC | MEDIUM TYPE: Floppy disk | | |
| CC | COMPUTER: IBM PC compatible | | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | |
| CC | SOFTWARE: Patentin Release #1.0, Version #1.25 | | |
| CC | CURRENT APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US/08/597,545 | | |
| CC | FILING DATE: | | |
| CC | CLASSIFICATION: 435 | | |
| CC | PRIOR APPLICATION DATA: | | |
| CC | APPLICATION NUMBER: US 07/989,537 | | |
| CC | FILING DATE: 11-DEC-1992 | | |
| CC | ATTORNEY/AGENT INFORMATION: | | |
| CC | NAME: BENT, Stephen A. | | |
| CC | REGISTRATION NUMBER: 29,768 | | |
| CC | REFERENCE/DOCKET NUMBER: 40399/166 NIHID | | |
| CC | TELECOMMUNICATION INFORMATION: | | |
| CC | TELEPHONE: (202)672-5300 | | |
| CC | TELEFAX: (202)672-5399 | | |
| CC | TELEX: 904136 | | |
| CC | INFORMATION FOR SRO ID NO: 3: | | |
| CC | SEQUENCE CHARACTERISTICS: | | |
| CC | LENGTH: 1556 base pairs | | |
| CC | TYPE: nucleic acid | | |
| CC | STRANDEDNESS: double | | |
| CC | TOPOLOGY: linear | | |
| CC | ORIGINAL SOURCE: | | |
| CC | STRAIN: Human Dlx | | |
| Q9 | Sequence 1556 BP; 302 A; 514 C; 440 G; 300 T; 0 other; | | |

```

Query Match      1.3%: Score 37; DB 1; Length 1556;
Best Local Similarity 57.5%; Pred. No. 7.35e-09;
Matches 199; Conservative 0; Mismatches 144; Indels 3; Gaps 2

Db    271 AAAAAAGATTCGACAGATGCATCATTTTCTCAGGTGCCATGTCGGCTGGCGAGGTCC 330
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    848 AACATGAGCATTTGGACAACAACCATTCAATCATCGTTTGCAACTGGGCTGGAAAGGAGCT 907
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    331 TTTGTGACAGTGGCTGACCTCTCTCCGGCTGCTTACGCACTGTGGAGAACC CGG 390
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    908 TTGTGAAGAGTAGCTTTGTGNAACCGAAGCTCATTCATGAGCACCTGCAACAAACCTGGA 967
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    391 AGTGAATTTGACGACGAGGCTGGGACGGGAGGACTGTGATAGAGATGTTGGGCTGCT 450
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    968 CTGTGATCTGCACAGAGGATTTGGGAGGCTTTGACTGCACACAGATCTGAACTAGTA 1027
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    451 CCT-CGGCC--CCTGTGCCAACAGAGGAGCTGCTGAGCCTGAGAGTGGCCCTATG 507
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    1028 CCACACACAGACCTTCGANAAGATGGCGGAGACCTGCTTCAACCGGGAGGAGATTGACA 1087
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    508 AATGCTCCTGTGCCCCCGGTTACTGGGAAAAGACGTGCCAGAAAAGAGAGGCGCTGTG 567
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    1088 CATGCAAAATGCCCCTCAGAGATAACGTGTGATGATTTGGAAAAAGATCTACTCTGCG 1147
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    568 TGAATGACGGCTCCCTGCGACGACGAGGACACTGGGTGATGA 613
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    1148 ATGCCGATGTCAATCCTGCGCAAGATGGTGTACTGATGATGA 1193
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT      10
ID          US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC          xxxxxx
DT          01-JAN-1900
DE          Sequence 14, Application US/08232463.
CC          Sequence 14, Application US/08232463
CC          Patent No. 5670367
CC          GENERAL INFORMATION:
CC          APPLICANT: DORNER, F.
CC          APPLICANT: SCHEFFELINGER, F.
CC          APPLICANT: PALMER, F. G.
CC          TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC          NUMBER OF SEQUENCES: 52
CC          CORRESPONDENCE ADDRESS:
CC          ADDRESSEE: Foley & Lardner
CC          STREET: 1800 Diagonal Road, Suite 500
CC          CITY: Alexandria
CC          STATE: VA
CC          COUNTRY: USA
CC          ZIP: 22313-0299
CC          COMPUTER READABLE FORM:
CC          MEDIUM TYPE: Floppy disk
CC          COMPUTER: IBM PC compatible
CC          OPERATING SYSTEM: PC-DOS/MS-DOS
CC          SOFTWARE: Patentin Release #1.0, Version #1.25
CC          CURRENT APPLICATION DATA:
CC          APPLICATION NUMBER: US/08/232,463
CC          FILING DATE:
CC          CLASSIFICATION: 435
CC          PRIOR APPLICATION DATA:
CC          APPLICATION NUMBER: US/07/935,313
CC          FILING DATE:
CC          APPLICATION NUMBER: EP 91 114 300.6
CC          FILING DATE: 26-AUG-1991
CC          ATTORNEY/AGENT INFORMATION:
CC          NAME: BENT, Stephen A.
CC          REGISTRATION NUMBER: 29,768
CC          REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC          TELECOMMUNICATION INFORMATION:
CC          TELEPHONE: (703)836-9300
CC          TELEFAX: (703)683-4109
CC          TELEX: 899149
CC          INFORMATION FOR SEQ ID NO: 14:

```

```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PT29pt-F1s
SQ Sequence 7218 BP, 1944 A; 1491 C; 1486 G; 1929 T; 368 other;

Query Match 1.38; Score 38; DB 1; Length 7218;
Best Local Similarity 2.08; Pred. No. 1.74e-09;
Matches 4; Conservative 116; Mismatches 82; Indels 0; Gaps 0;

Db 1062 GCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1121
QY 330 GGGTTTCGGCTGCTAAAGCACTACCAAGCCACCATCGACACCACTCCAGTGCC 389
Db 1122 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1181
QY 390 CTACGGGACGTGATCACGCCCATCTCGCGAGAACTCGGTCAATCTGACCAAGCCCA 449
Db 1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241
QY 450 GCGCTTCCAGAACAGGCTTCCAGATCCATCCAGTTCCTCTCGTTCTCATGGCC 509
Db 1242 YYYYYYYYYYYYYYYYYYYYYYYY 1263
QY 510 GGGTACCTTCTCGTGATCGTC 531

RESULT 11
ID PCT-US96-02331-14 STANDARD; DNA; UNC; 4255 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application PC/TUS9602331.
CC Sequence 14, Application PC/TUS9602331
CC GENERAL INFORMATION:
CC APPLICANT: The Board of Trustees of the Leland Stanford Junior
CC APPLICANT: University
CC APPLICANT: Board of Regents, The University of Texas System
CC TITLE OF INVENTION: Methods and Compositions for Altering
CC TITLE OF INVENTION: Sexual Behavior
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/02331
CC FILING DATE: 09-FEB-1996
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/386,495
CC FILING DATE: 10-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles K.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 8600-0153.41
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4255 base pairs
CC TYPE: nucleic acid

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: fruitless transcript in Fig. 7E
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1507..4032
SQ Sequence 4255 BP; 1273 A; 1130 C; 1000 G; 852 T; 0 other;
Matches 35; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query Match 1.18; Score 32; DB 2; Length 4255;
Best Local Similarity 92.18; Pred. No. 8.10e-06;
Matches 35; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 2073 CTCGGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2110
QY 2226 CTCAGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2263

RESULT 12
ID PCT-US96-02331-9 STANDARD; DNA; UNC; 4835 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 9, Application PC/TUS9602331.
CC Sequence 9, Application PC/TUS9602331
CC GENERAL INFORMATION:
CC APPLICANT: The Board of Trustees of the Leland Stanford Junior
CC APPLICANT: University
CC APPLICANT: Board of Regents, The University of Texas System
CC TITLE OF INVENTION: Methods and Compositions for Altering
CC TITLE OF INVENTION: Sexual Behavior
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/02331
CC FILING DATE: 09-FEB-1996
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/386,495
CC FILING DATE: 10-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles K.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 8600-0153.41
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4835 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: Fru#1 CDNA
CC FEATURE:
CC NAME/KEY: CDS
```



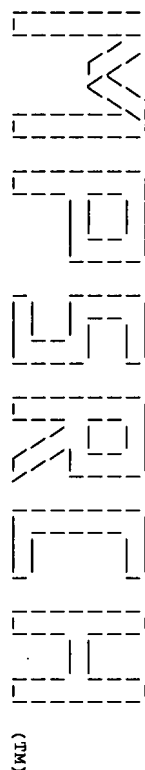
```

CC Sequence 1, Application US/08385500
CC Patent No. 5712117
CC GENERAL INFORMATION:
CC APPLICANT: Sprecher, Cindy A.
CC TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
CC TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/385,500
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parmelee, Steven W.
CC REGISTRATION NUMBER: 31,990
CC REFERENCE/DOCKET NUMBER: 13952-21
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 467-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1425 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 92..1213
CC OTHER INFORMATION: /product= "CYTOPLASMIC
CC OTHER INFORMATION: ANTIPROTEINASE-2 PROTEIN"
CC
CC SQ Sequence 1425 BP; 437 A; 301 C; 357 G; 330 T; 0 other;

Query Match 1.0%; Score 28; DB 1; Length 1425;
Best Local Similarity 91.2%; Pred. No. 1.67e-03;
Matches 31; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 37 CAGCAGCGCGCGCGCGCGCGCGCAGCAGCAGC 70
      ||||| || ||||| ||||| ||||| ||
QY 2228 CAGCAGCAGCAGCGCGCGCGCGCGCAGCAGC 2261

```



(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jun 13 08:37:40 1998; Maspar time 2790.42 Seconds
1384.359 Million cell updates/sec

Tabular output not generated.

Title: >US-08-083-590A-1
Description: (1-2892) from US08083590A.seq
Perfect Score: 2892
N.A. Sequence: 1 GAATTCGAGAGATATATCA.....GCCGTTGCATCCGATTC 2892
Comp: CTTAAGCTCTCTTAATAGT.....CGGCAAGCTTAGGCTTAG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est54
1:em_est1 2:em_est3
Database: genbank-est106
3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
16:gb_est21 17:gb_est22 18:gb_est3 19:gb_est4 20:gb_est5
21:gb_est6 22:gb_est7 23:gb_est8 24:gb_est9 25:gb_ests
26:gb_ests

Statistics: Mean 11.911; Variance 2.337; scale 5.097

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-------------|----------------------------------|
| 1 | 436 | 15.1 | 442 | 13 | AA538738 | LD18334.5prime LD Dros 0.00e+00 |
| 2 | 377 | 13.0 | 379 | 13 | AA540186 | LD19584.5prime LD Dros 0.00e+00 |
| 3 | 150 | 5.2 | 568 | 12 | AA142228 | CK00091.3prime CK Dros 2.05e-235 |
| 4 | 78 | 2.7 | 247 | 23 | AA015264 | mh30d12.r1 Soares mous 5.84e-93 |
| 5 | 60 | 2.1 | 252 | 13 | AA754459 | 97SN1787 Rice Immature 6.76e-60 |
| 6 | 57 | 2.0 | 247 | 13 | AA754458 | 97SN1787 Rice Immature 1.40e-54 |
| 7 | 59 | 2.0 | 252 | 13 | AA754459 | 97SN1787 Rice Immature 4.07e-58 |
| 8 | 55 | 1.9 | 570 | 21 | AA051023 | mg73f08.r1 Soares mous 4.48e-51 |
| 9 | 51 | 1.8 | 247 | 13 | AA051023 | 97SN1784 Rice Immature 3.60e-44 |
| 10 | 51 | 1.8 | 602 | 23 | AA046860 | zf14h11.r1 Soares feta 3.60e-44 |
| 11 | 50 | 1.7 | 445 | 10 | AA619107 | yf66f10.r1 Homo sapien 1.82e-42 |
| 12 | 42 | 1.5 | 445 | 10 | AA619107 | vo68b02.r1 Soares mous 2.96e-29 |
| 13 | 42 | 1.5 | 485 | 23 | AA024229 | mh99e03.r1 Soares mous 2.96e-29 |

| | | | | | | |
|----|----|-----|------|----|----------|---------------------------------|
| 14 | 37 | 1.3 | 589 | 24 | AA238559 | my35e01.r1 Barstead mo 1.89e-21 |
| 15 | 39 | 1.3 | 592 | 23 | AA153148 | mr47c02.r1 Stratagene 1.60e-24 |
| 16 | 39 | 1.3 | 642 | 21 | AA487442 | mj44c04.r1 Soares mous 1.60e-24 |
| 17 | 39 | 1.3 | 766 | 23 | AA165945 | ms50g09.r1 Life Tech m 1.60e-24 |
| 18 | 39 | 1.3 | 2275 | 12 | AF034173 | Homo sapiens ntcon2 co 1.60e-24 |
| 19 | 36 | 1.2 | 378 | 12 | AA141415 | CK01709.5prime CK Dros 6.08e-70 |
| 20 | 36 | 1.2 | 390 | 10 | W13561 | ma85d08.r1 Soares mous 5.51e-17 |
| 21 | 34 | 1.2 | 527 | 23 | AA171296 | ms57a11.r1 Life Tech m 1.87e-18 |
| 22 | 35 | 1.2 | 786 | 23 | AA171309 | ms57e11.r1 Life Tech m 1.87e-18 |
| 23 | 35 | 1.2 | 2275 | 12 | AF034173 | Homo sapiens ntcon2 co 1.87e-18 |
| 24 | 32 | 1.1 | 254 | 14 | R50365 | yj59e07.s1 Homo sapien 4.08e-14 |
| 25 | 32 | 1.1 | 269 | 22 | AA059314 | zf66b05.s1 Soares retl 4.08e-14 |
| 26 | 33 | 1.1 | 328 | 18 | H39899 | yp02h01.s1 Homo sapien 1.54e-15 |
| 27 | 31 | 1.1 | 391 | 11 | AA683105 | ae81h06.s1 Stratagene 1.02e-12 |
| 28 | 32 | 1.1 | 416 | 14 | R72432 | yj99f05.s1 Homo sapien 4.08e-14 |
| 29 | 31 | 1.1 | 422 | 14 | R50026 | yj59e07.r1 Homo sapien 1.02e-12 |
| 30 | 32 | 1.1 | 459 | 14 | RA6751 | yj54a01.s1 Homo sapien 4.08e-14 |
| 31 | 31 | 1.1 | 470 | 8 | AA215625 | zr98c02.s1 NCI CGAP GC 1.02e-12 |
| 32 | 31 | 1.1 | 518 | 12 | AA246536 | LD05485.5prime LD Dros 4.08e-14 |
| 33 | 32 | 1.1 | 607 | 25 | B16636 | 342p20.TPB Ctr978SKA1 4.08e-14 |
| 34 | 33 | 1.1 | 729 | 12 | HSU69180 | Human DRES 30 mRNA seq 1.54e-15 |
| 35 | 33 | 1.1 | 136 | 24 | AA139263 | mq02e06.r1 Soares feta 2.40e-11 |
| 36 | 36 | 1.0 | 297 | 22 | W67931 | zd38g09.r1 Soares feta 5.30e-10 |
| 37 | 30 | 1.0 | 337 | 23 | AA020103 | mh49e03.r1 Soares mous 2.40e-11 |
| 38 | 30 | 1.0 | 346 | 19 | N24541 | yx75c06.s1 Homo sapien 2.40e-11 |
| 39 | 30 | 1.0 | 353 | 26 | HS107XC7 | H. sapiens (D11S906) D 2.40e-11 |
| 40 | 30 | 1.0 | 390 | 3 | H49928 | ya99d07.r1 Homo sapien 2.40e-11 |
| 41 | 29 | 1.0 | 414 | 18 | H16136 | yl28h01.r1 Homo sapien 5.30e-10 |
| 42 | 29 | 1.0 | 421 | 12 | AA733056 | zy79b12.s1 Soares feta 5.30e-10 |
| 43 | 29 | 1.0 | 444 | 20 | W16654 | zb06h12.r1 Soares feta 5.30e-10 |
| 44 | 29 | 1.0 | 457 | 18 | H20919 | yn65b09.r1 Homo sapien 5.30e-10 |
| 45 | 29 | 1.0 | 487 | 18 | H42553 | yo63b10.r1 Homo sapien 5.30e-10 |

ALIGNMENTS

RESULT 1
LOCUS AA538738 442 bp mRNA
DEFINITION LD18334.5prime LD Drosophila melanogaster emb106 BlueScript
Drosophila melanogaster cDNA clone LD18334 5prime similar to
X05140: Drosophila melanogaster mRNA fragment of D1 locus with
EGF-like repeats (EGF-epidermal growth factor); Y00222:
D.melanogaster D1 mRNA for delta protein, involved in neurogenesis,
mRNA sequence.

ACCESSION AA538738
NID g2285254
KEYWORDS EGF.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

Eukaryote; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
1 (bases 1 to 442)

REFERENCE Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
Broksfeld, P., Lewis, S. and Rubin, G.M.
BDGP/HMT Drosophila EST Project
Unpublished (1997)

TITLE

JOURNAL

COMMENT

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 USA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://fruitfly.berkeley.edu/EST>, estefruitfly.berkeley.edu
Plate: 183 row: C column: 10
High quality sequence stop: 390.

FEATURES

source

Location/Qualifiers
1..442
/organism="Drosophila melanogaster"
/note="Organ: embryo; Vector: Bluescript SK. site: 1;
ECORI. site: 2: XhoI; Constructed using Strategene ZAP-cDNA
Synthesis Kit. Oligo dt-primed and directionally cloned at
ECORI and XhoI in Bluescript SK(+/-)"

/db_xref="BDGP_EST:BDcln017469"
/db_xref="taxon:7227"
/clone="LD18334"
/clone_lib="LD Drosophila melanogaster embryo BlueScript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"

BASE COUNT 110 a 117 c 127 g 88 t

ORIGIN

Query Match 15.1%; Score 436; DB 13; Length 442;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 439; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 1 CCAGATGTTGTACTGATGATGAGCGGCACAAACCGGTACAGTGTCTATTG 60
|||||
QY 1167 CCAGATGTTGTACTGATGATGAGCGGCACAAACCGGTACAGTGTCTATTG 1226
|||||
Db 61 CGCCAAACGCTGGAGGGAAGATGTGCGAGGAGAAAGTCTCACGTGTTGCGACAAACC 120
|||||
QY 1227 CGCCAAACGCTGGAGGGAAGATGTGCGAGGAGAAAGTCTCACGTGTTGCGACAAACC 1286
|||||
Db 121 CTGTATCAGGAATCTGCCCAACGTTCTCTGTTTGGGAAGCAAGGTCAGGGCTA 180
|||||
QY 1287 CTGTATCAGGAATCTGCCCAACGTTCTCTGTTTGGGAAGCAAGGTCAGGGCTA 1346
|||||
Db 181 CCAATGCGAATGTCATTTGGCTACAGCGGACCAACTGGGATCTCCAGTGGACAACTG 240
|||||
QY 1347 CCAATGCGAATGTCATTTGGCTACAGCGGACCAACTGGGATCTCCAGTGGACAACTG 1406
|||||
Db 241 CAGTCCGAATCTCTGATTAACGTTGGAAGTGTCTACGCGAGGGAAGTGTATTGGCCC 300
|||||
QY 1407 CAGTCCGAATCTCTGATTAACGTTGGAAGTGTCTACGCGAGGGAAGTGTATTGGCCC 1466
|||||
Db 301 ACGGGATTTTCGGGAACGAGATGCGAGACCAACATTTGACGATTTGTTGGCCACCACTG 360
|||||
QY 1467 ACGGGATTTTCGGGAACGAGATGCGAGACCAACATTTGACGATTTGTTGGCCACCACTG 1526
|||||
Db 361 CAGAACGAGGACCTGATAGATATGTTCAACCAATATCGTGGCAATGCTTCCCGG 420
|||||
QY 1527 CAGAACGAGGACCTGATAGATATGTTCAACCAATATCGTGGCAATGCTTCCCGG 1586
|||||
Db 421 TTTCCATGCGACCCACTGTAGT 442
|||||
QY 1587 TTTCCATGCGACCCACTGTAGT 1608
|||||

RESULT 2 AA540186 379 bp mRNA EST 19-JAN-1998
LOCUS LD19584.5prime LD Drosophila melanogaster embryo BlueScript
DEFINITION Drosophila melanogaster cDNA clone LD19584.5prime similar to
L03725: Drosophila sp. epidermal growth factor-like element (Notch
and Delta) gene regions; Y00222: D.melanogaster D1 mRNA for delta
protein, involved in neurogenesis, mRNA sequence.

ACCESSION AA540186
NID 92286620
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryotae; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 379)
AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (1997)
COMMENT

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947

Email: http://fruitfly.berkeley.edu/EST_est@fruitfly.berkeley.edu
Plate: 195 row: 6 column: 12
High quality sequence stop: 182.

FEATURES

source

1. 379
Location/Qualifiers
/organism="Drosophila melanogaster"
/note="Organ: embryo; Vector: BlueScript SK; Site: 1;
EcoRI; Site: 2; XhoI; Constructed using Stratagene ZAP-cDNA
Synthesis kit. Oligo dr-primed and directionally cloned at
EcoRI and XhoI in BlueScript SK(+/-)"
/db_xref="BDGP_EST:BDcln018671"
/db_xref="taxon:7227"
/clone="LD19584"
/clone_lib="LD Drosophila melanogaster embryo BlueScript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"

BASE COUNT 97 a 105 c 103 g 74 t

ORIGIN

Query Match 13.0%; Score 377; DB 13; Length 379;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 CCAAGTTCTGCGCGCCCGCGACGATTCTTTGGACACTCGACTTGTCTGGAGACGGCG 60
|||||
QY 722 CCAAGTTCTGCGCGCCCGCGACGATTCTTTGGACACTCGACTTGTCTGGAGACGGCG 781
|||||
Db 61 AAATATCTGTTTACCAGGATGCGAGGGGATTACTGTACATATCCCAATGCGCCAAAG 120
|||||
QY 782 AAATATCTGTTTACCAGGATGCGAGGGGATTACTGTACATATCCCAATGCGCCAAAG 841
|||||
Db 121 GCTGTGAACATGACATTTGGACAGCCCAATCAATCGTTCGCCAACCTGGGCTGGAAG 180
|||||
QY 842 GCTGTGAACATGACATTTGGACAGCCCAATCAATCGTTCGCCAACCTGGGCTGGAAG 901
|||||
Db 181 GAGCCTTTGTCACAGAGTGTCTGGAACCGCAACTGTCATCCATGCGACCTGCAACAAAC 240
|||||
QY 902 GAGCCTTTGTCACAGAGTGTCTGGAACCGCAACTGTCATCCATGCGACCTGCAACAAAC 961
|||||
Db 241 CTTGGACTTGCATCTGCAACGAGGTTGGGGAGGCTTGTACTGCAACCCAGGATCTGAACT 300
|||||
QY 962 CTTGGACTTGCATCTGCAACGAGGTTGGGGAGGCTTGTACTGCAACCCAGGATCTGAACT 1021
|||||
Db 301 ACTGCACCAACACACACCTGCAAGAAATGGCGGAGACCTGCTTCACACCGCGGAGGAT 360
|||||
QY 1022 ACTGCACCAACACACACCTGCAAGAAATGGCGGAGACCTGCTTCACACCGCGGAGGAT 1081
|||||
Db 361 TGTACACATGCAAAATGCGC 379
|||||
QY 1082 TGTACACATGCAAAATGCGC 1100
|||||

RESULT 3

LOCUS

DEFINITION

ACCESSION

NID

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA142228 568 bp mRNA EST 12-JAN-1998
CK00091.3prime CK Drosophila melanogaster embryo BlueScript
Drosophila melanogaster cDNA clone CK00091.3prime similar to 0;
mRNA sequence.
AA142228
q1703880
EST.
fruit fly.
Drosophila melanogaster
Eukaryotae; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
1 (bases 1 to 568)
Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
Other ESTs: CK00091.5prime.contig
Contact: Harvey, D.

JOURNAL Unpublished (1998)


```
/sbin= C3/B5/00
/note="Vector: PT7T3D-Pac (Pharmacia) with a modified
```

seq primer: 5' reverse primer:
location/qualifiers

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:369898

Seq primer: -28m3 rev1 ET from Amersham
High quality sequence stop: 381.

FEATURES

source

1. 592

Location/Qualifiers

/organism="Mus musculus"

/strain="NIH/Swiss"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
CAATCGGCACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3'"

/db_xref="taxon:10090"

/clone="604466"

/clone.lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/dev_stage="13 day embryos"

/lab_host="SOLR (kanamycin resistant)"

<1. .>592

mRNA

BASE COUNT

106 a 196 c 176 g 114 t

ORIGIN

Query Match

1.38; Score 39; DB 23; Length 592;

Best Local Similarity 63.84; Pred. No. 1.60e-24;

Matches 90; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

| | | | |
|----|-----|----------------------------------------------------------|-----|
| DB | 239 | ATGATCTGCGGCTGACATGTGCGAGTGGCCATGTTGGTGGAGGATCCCTCT | 298 |
| OY | 851 | ATGGACATTCGCGACAAACCAATCAATGCGTTGCCACTGGGCTGGAAGGAGCCTGT | 910 |
| DB | 299 | GTGACAGTGTGTACTGCGCTGCTGTGTCATGAGTCTGCAAGAACCATGGCAGT | 358 |
| OY | 911 | GCAACGAGTGCCTGTGGAACCGAATGCATCATGCGACCTGCACAAACCTGGACTT | 970 |
| DB | 359 | GCATCTGCAAGATGGCTGGG | 379 |
| OY | 971 | GCATCTGCAAGAGGCTGGG | 991 |

Search completed: Sat Jun 13 09:56:28 1998
Job time : 4728 secs.

118184969 CA: 118(19)184969g JOURNAL
 Mouse mammary tumor gene int-3: a member of the notch gene family
 transforms mammary epithelial cells
 AUTHOR(S): Robbins, Joan; Blondel, Bruno J.; Gallahan, Daniel; Callahan, Robert
 LOCATION: Lab. Tumor Immunol. Biol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
 JOURNAL: J. Virol. DATE: 1992 VOLUME: 66 NUMBER: 4 PAGES: 2594-9
 CODEN: JOVIAM ISSN: 0022-538X LANGUAGE: English
 SECTION:
 CA203004 Biochemical Genetics
 CA213XXX Mammalian Biochemistry
 IDENTIFIERS: mouse mammary tumor gene int3 transformation, epithelium transformation mouse notch like gene
 DESCRIPTORS:
 Epithelium...
 cell, of mammary gland of mouse, gene int-3 transformation of
 Deoxyribonucleic acid sequences...
 for gene int-3 protein of mouse mammary tumor
 Proteins, specific or class, gene int-3...
 gene for, of mouse mammary tumor, sequence of and epithelial cell transformation by
 Virus, animal, murine mammary tumor...
 gene int-3 integration site for, sequence of and epithelial cell transformation by
 Mouse...
 gene int-3 of mammary tumor of, sequence of and epithelial cell transformation by
 Gene, animal, Notch...
 mouse mammary tumor gene int-3 like, of Drosophila melanogaster, sequence of and epithelial cell transformation by
 Mammary gland, neoplasm...
 notch family gene int-3 of mouse, sequence of and epithelial cell transformation by
 Drosophila melanogaster...
 notch gene of, mouse mammary tumor gene int-3 as member of family of, sequence of and epithelial cell transformation by
 Protein sequences...
 of gene int-3 protein, of mouse mammary tumor
 Gene, animal, int-3...
 of mouse mammary tumor, sequence of and epithelial cell transformation by
 Deoxyribonucleic acids, repetitive...
 Saccharomyces cerevisiae cell cycle gene cdc-10 homolog, in gene int-3 of mouse mammary tumor
 CAS REGISTRY NUMBERS:
 146991-60-8 amino acid sequence of, complete
 139861-79-3 nucleotide sequence of

16/7/176 (Item 3 from file: 399)
 DIALOG(R) File 399:CA SEARCH(R)
 (c) 1996 American Chemical Society. All rts. reserv.

120130975 CA: 120(11)130975d DISSERTATION
 TAN-1, the human homolog of Drosophila "Notch", is involved in chromosomal translocations in human lymphoblastic neoplasma

AUTHOR(S): Ellisen, Leif William
LOCATION: Stanford Univ., Stanford, CA, USA
DATE: 1992 PAGES: 80 pp. CODEN: DABBBB LANGUAGE: English CITATION:
Diss. Abstr. Int. B 1993, 53(7), 3307 AVAIL: Univ. Microfilms Int., Order
No. DA9234046

SECTION:

CA214001 Mammalian Pathological Biochemistry

CA203XXX Biochemical Genetics

IDENTIFIERS: gene TAN1 chromosome translocation lymphoblastic leukemia

DESCRIPTORS:

Leukemia, T-cell acute lymphocytic...

gene TAN-1 translocation in, in human

Recombination, genetic, translocation...

of gene TAN-1, in human T-cell acute lymphoblastic leukemia

Gene, animal...

TAN-1, chromosomal translocation of, in human T-cell lymphoblastic
leukemia

Chromosome, human 7...

TAN-1 gene translocation from, in human T-cell lymphoblastic leukemia

16/7/141 (Item 28 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

06701790 89003790

Expression of the differentiation antigen F7D6 in tumorous tissues of *Drosophila*.

Bedian V; Jungklaus CE

Department of Biology, Clarkson University, Potsdam, New York 13676.

Dev Genet (UNITED STATES) 1987, 8 (3) p165-77, ISSN 0192-253X

Journal Code: DEG

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The 63-kDa antigen recognized by the monoclonal antibody F7D6 is present in all *Drosophila* embryonic cells and disappears from most tissues as each one reaches its final, differentiated state. Larval tissues lose the antigen around the time of hatching, imaginal tissues lose it during metamorphosis, and germ cells lose it during gametogenesis (Bedian et al: Devel Biol 115:105-118, 1986). The nervous system and spontaneously contracting musculature of the gut and gonads are exceptions and remain antigen positive at all stages. The F7D6 antigen appears to be associated with dividing, undifferentiated cells and electrogenic cells. This prompted us to test tumors for antigen presence. We tested four different recessive mutants that give rise to four different types of tumorous transformation: the embryonic tumor Notch, several larval melanotic tumors, the imaginal disc tumor 1(2)gl, and three alleles of the ovarian tumor otu. In all cases, tumorous tissues in homozygotes contained the F7D6 antigen. The electrophoretic mobility of the antigen appeared to be unaltered in tumorous tissues compared to normal cells, but the antigen is expressed at higher levels. The antigen is found on the cytoplasmic surface of plasma membranes and appears to be a marker of undifferentiated normal and tumorous cells. Similarities and differences between the F7D6 antigen and *Drosophila* c-src protein are discussed.

16/7/139 (Item 26 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

06969982 89271982

[Adult T-cell leukemia (ATL)]

Takatsuki K

Kumamoto University Medical School.

Gan To Kagaku Ryoho (JAPAN) Apr 1989, 16 (4 Pt 2-1) p880-5, ISSN

0385-0684 Journal Code: 6T8

Languages: JAPANESE Summary Languages: ENGLISH

Document type: JOURNAL ARTICLE English Abstract

Adult T-cell leukemia (ATL) was first discovered and reported in Japan, where it has a high incidence in the southwest region. The retrovirus HTLV-I (human T-cell lymphotropic virus type I) is considered to be related to its etiology. In ATL endemic areas, HTLV-I carriers are found at a fairly high percentage even among healthy individuals. ATL shows diverse clinical features. It can be divided into 5 types (acute type, chronic type, smoldering type, crisis type, and lymphoma type). ATL cells originate from the CD4-positive subset of peripheral T cells; they show a characteristic notch in the nucleus and a lobulation tendency. ATL resists chemotherapy, and patients with acute and lymphoma types have quite a poor

prognosis. A definite diagnosis of ATL is made by documenting the presence of HTLV-I proviral DNA in the DNA of tumor cells. HTLV-I infection is caused by transmission of live lymphocytes via three routes (from mother to children, from males to females, and by transfusion). Familial occurrence of ATL is frequently seen. HTLV-I infection is also seen in other countries, but its incidence is highest in Japan. It is thus an urgent task for Japanese physicians to eliminate HTLV-I infection.

16/7/128 (Item 15 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

08122975 92260975
Cancer, chromosomes, and genes.
Nowell PC
Department of Pathology and Laboratory Medicine, University of Pennsylvania, School of Medicine, Philadelphia.
Lab Invest (UNITED STATES) Apr 1992, 66 (4) p407-17, ISSN 0023-6837
Journal Code: KZ4
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, ACADEMIC
(67 Refs.)

16/7/123 (Item 10 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09041700 94356700
Analysis of phenotypic abnormalities and cell fate changes caused by dominant activated and dominant negative forms of the Notch receptor in Drosophila development.
Rebay I; Fortini ME; Artavanis-Tsakonas S
Howard Hughes Medical Institute, Department of Cell Biology, Yale University, New Haven, Connecticut 06536-0812.
C R Acad Sci III (FRANCE) Sep 1993, 316 (9) p1097-123, ISSN 0764-4469
Journal Code: CA1
Languages: ENGLISH, FRENCH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, TUTORIAL
The Notch gene of Drosophila plays an important role in cell fate specification throughout development. The Notch protein contains a large extracellular domain of 36 EGF-like repeats as well as 3 Notch/lin-12 repeats and an intracellular domain with 6 cdc10/ankyrin repeats, motifs which are highly conserved in several vertebrate Notch homologues [1-7]. In this review we summarize the results of two recent studies which attempt to establish structure-function relationships of the various domains of the Notch gene product [8, 9]. The functions of various structural domains of the Notch protein in vivo were investigated using a series of deletion mutants which have been ectopically expressed either under the hsp70 heat-shock promoter or under the sevenless eye-specific promoter. Truncation of the extracellular domain of Drosophila Notch produces an activated receptor as judged by its ability to cause phenotypes matching those of gain-of-function alleles or duplications of the Notch locus [8]. Equivalent truncations of vertebrate Notch-related proteins have been associated with malignant neoplasms and other developmental abnormalities [3, 6, 10, 11]. In contrast, dominant negative phenotypes result from overexpression of a protein lacking most intracellular sequences. These

results were extended by an analysis of activated Notch function at single-cell resolution in the *Drosophila* compound eye [9]. It was shown that while overexpression of full-length Notch in defined cell types has no apparent effects, overexpression of activated Notch in the same cells transiently blocks their proper cell-fate commitment, causing them to either adopt incorrect cell fates or to differentiate incompletely. Moreover, an activated Notch protein lacking the transmembrane domain is translocated to the nucleus, raising the possibility that Notch may participate directly in nuclear events. (68 Refs.)

16/7/116 (Item 3 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09600265 96121865
Oncogene activation and oncogene cooperation in MMTV-induced mouse mammary cancer.

van Leeuwen F; Nusse R
Howard Hughes Medical Institute, Department of Developmental Biology, Beckman Center, Stanford University, Medical Center, CA 94305-5428, USA.
Semin Cancer Biol (UNITED STATES) Jun 1995, 6 (3) p127-33, ISSN 1044-579X Journal Code: A6Y

Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, TUTORIAL
In this short review, we will give an overview of the various genes that can be activated by insertion of proviral DNA of the mouse mammary tumor virus during the formation of mammary cancer. These genes fall within three families: Wnt genes, FGF genes and Notch-related genes. We will summarize our current understanding of the roles of these genes in tumorigenesis and in normal development, and the mechanisms of action of their gene products. Finally, we will give some examples of cooperation between these genes in various biological settings. (91 Refs.)

16/7/113 (Item 6 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

04166645 PASCAL No.: 75-0002761
DEVELOPMENTAL CAPACITIES OF BENIGN AND MALIGNANT NEOPLASMS OF DROSOPHILA
GATEFF E; SCHNEIDERMAN H A
CENT. PATHOBIOL., UNIV. CALIFORNIA, IRVINE, CALIF. 92664, USA
Journal: ARCH. (WILHELM ROUX) ENTWICKL.-MECH. ORG., 1974, 176 (1) 23-65
Availability: CNRS-3573
No. of Refs.: 4P.1/2
Document Type: P (SERIAL) ; A (ANALYTIC)
Country of Publication: FEDERAL REPUBLIC OF GERMANY
Language: ENGLISH Summary Language: GERMAN

16/7/112 (Item 5 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

10278296 PASCAL No.: 92-0484208
Expression pattern of Motch, a mouse homolog of *Drosophila* notch, suggests an important role in early postimplantation mouse development

DEL AMO F F; SMITH D E; SWIATEK P J; GENDRON-MAGUIRE M; GREENSPAN R J;
MCMAHON A P; GRIDLEY T
Roche inst. molecular biology, dep. cell developmental biology, Nutley NJ
07110, USA

Journal: Development : (Cambridge), 1992, 115 (3) 737-744
ISSN: 0950-1991 Availability: INIST-7560; 354000020190410090

No. of Refs.: 1 p.1/2

Document Type: P (Serial) ; A (Analytic)

Country of Publication: United Kingdom

Language: English

The Notch gene of Drosophila encodes a large transmembrane protein involved in cell-cell interactions and cell fate decisions in the Drosophila embryo. To determine if a gene homologous to Drosophila Notch plays a role in early mouse development, we screened a mouse embryo cDNA library with probes from the Xenopus Notch homolog, Xotch. A partial cDNA clone encoding the mouse Notch homolog, which we have termed Motch, was used to analyze expression of the Motch gene. Motch transcripts were detected in a wide variety of adult tissues, which included derivatives of all three germ layers

16/7/110 (Item 3 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

12471360 PASCAL No.: 96-0134151

Epithelial expression and chromosomal location of human TLE genes :
implications for notch signaling and neoplasia

YANLING LIU; GHASSAN DEHNI; PURCELL K J; SOKOLOW J; CARCANGIU M L;
ARTAVANIS-TSAKONAS S; STIFANI S

McGill univ., Montreal neurological inst., Montreal PQ H3A 2B4, Canada

Journal: Genomics : (San Diego, CA), 1996, 31 (1) 58-64

ISSN: 0888-7543 Availability: INIST-21389; 354000052959520080

No. of Refs.: 26 ref.

Document Type: P (Serial) ; A (Analytic)

Country of Publication: USA

Language: English

The TLE genes are the human homologues of Drosophila groucho, a member of the Notch signaling pathway. This pathway controls a number of different cell-fate choices in invertebrates and vertebrates. We are interested in investigating the functions of the TLE gene family during epithelial determination and carcinogenesis. We show that expression of individual TLE genes correlates with immature epithelial cells that are progressing toward their terminally differentiated state, suggesting a role during epithelial differentiation. In both normal tissues and conditions resulting from incorrect or incomplete maturation events, such as metaplastic and neoplastic transformations, TLE expression is elevated and coincides with Notch expression, implicating these molecules in the maintenance of the undifferentiated state in epithelial cells. We also show that TLE1 and TLE2 are organized in a tandem array at chromosomal location 19p13.3, while TLE3 maps to 15q22.

?t s16/7/2,3,6,9,10,11,12,14,15,17,23,25,40

>>>Format 7 is not valid in file 143

16/7/2 (Item 2 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

13041777 BIOSIS Number: 99041777

Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles

Pear W S; Aster J C; Scott M L; Hassserjian R P; Soffer B; Sklar J; Baltimore D

Massachusetts Inst. Technol., Room 68-380, 77 Massachusetts Ave., Cambridge, MA 02139-4307, USA

Journal of Experimental Medicine 183 (5). 1996. 2283-2291.

Full Journal Title: Journal of Experimental Medicine

ISSN: 0022-1007

Language: ENGLISH

Print Number: Biological Abstracts Vol. 102 Iss. 002 Ref. 023950

Notch is a highly conserved transmembrane protein that is involved in cell fate decisions and is found in organisms ranging from *Drosophila* to humans. A human homologue of Notch, TAN1, was initially identified at the chromosomal breakpoint of a subset of T-cell lymphoblastic leukemias/lymphomas containing a t(7;9) chromosomal translocation; however, its role in oncogenesis has been unclear. Using a bone marrow reconstitution assay with cells containing retrovirally transduced TAN1 alleles, we analyzed the oncogenic potential of both nuclear and extranuclear forms of truncated TAN1 in hematopoietic cells. Although the Moloney leukemia virus long terminal repeat drives expression in most hematopoietic cell types, retroviruses encoding either form of the TAN1 protein induced clonal leukemias of exclusively immature T cell phenotypes in approx 50% of transplanted animals. All tumors overexpressed truncated TAN1 of the size and subcellular localization predicted from the structure of the gene. These results show that TAN1 is an oncoprotein and suggest that truncation and overexpression are important determinants of transforming activity. Moreover, the murine tumors caused by TAN1 in the bone marrow transplant model are very similar to the TAN1-associated human tumors and suggest that TAN1 may be specifically oncotropic for T cells.

16/7/3 (Item 3 from file: 5)

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

(c) 1996 BIOSIS. All rts. reserv.

13034202 BIOSIS Number: 99034202

T cell leukemia-associated human notch-translocation-associated notch homologue has I-kappa-B-like activity and physically interacts with nuclear factor-kappa-B proteins in T cells

Guan E; Wang J; Laborda J; Norcross M; Baeuerle P A; Hoffman T

Lab. Cell Biol., Div. Monoclonal Antibodies, Cent. Biologics Eval. Res., Federal Drug Administration, Building 29B, Room 4G03, HFM558, 8800 Rockville Pike, Bethesda, MD 20892, USA

Journal of Experimental Medicine 183 (5). 1996. 2025-2032.

Full Journal Title: Journal of Experimental Medicine

ISSN: 0022-1007

Language: ENGLISH

Print Number: Biological Abstracts Vol. 102 Iss. 002 Ref. 016375

Translocation-associated Notch homologue (TAN-1), a gene originally cloned from the translocation breakpoint of a human T cell leukemia carrying a 9:7(q34.3) translocation, encodes a protein belonging to the Notch/Lin-12/Glp-1 receptor family. These receptors mediate the specification of numerous cell fates during development in invertebrates and vertebrates. The intracellular portion of Notch/TAN-1 contains six ankyrin repeats that are similar to those found in cytoplasmic I-kappa-B

proteins. I-kappa-B proteins are specific inhibitors of nuclear factor (NF)-kappa-B/Rel transcription factors. Here we show that TAN-1 has functional properties of an I-kappa-B-like regulator with specificity for the NF-kappa-B p50 subunit. A recombinant polypeptide corresponding to the cytoplasmic portion of TAN-1 (TAN-1-C) specifically inhibited the DNA binding of p50-containing NF-kappa-B complexes. When overexpressed in an appropriate cell line, TAN-1-C prevented kappa-B-dependent transactivation in transient reporter gene assays in a fashion similar to the structurally related protein, Bcl-3. TAN-1-C could activate kappa-B-dependent gene expression by attenuating the inhibitory effect of an excess of p50 homodimers. Immunoprecipitation experiments showed that the TAN-1 from a T cell line is associated with NF-kappa-B containing p50 and p65 subunits. These observations indicate that TAN-1-C may directly engage NF-kappa-B transcription factors and modulate nuclear gene expression.

16/7/6 (Item 6 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

12098006 BIOSIS Number: 98698006

Epithelial expression and chromosomal location of human TLE genes:
Implications for notch signaling and neoplasia

Liu Y; Dehni G; Purcell K J; Sokolow J; Carcangiu M L; Artavanis-Tsakonas S; Stifani S

Montreal Neurol. Inst., 3801 University St., Montreal, PQ H3A 2B4, Canada
Genomics 31 (1). 1996. 58-64.

Full Journal Title: Genomics

ISSN: 0888-7543

Language: ENGLISH

Print Number: Biological Abstracts Vol. 101 Iss. 007 Ref. 098287

The TLE genes are the human homologues of Drosophila groucho, a member of the Notch signaling pathway. This pathway controls a number of different gene family during epithelial determination and carcinogenesis. We show that expression of individual TLE genes correlates with immature epithelial cells that are progressing toward their terminally differentiated state, suggesting a role during epithelial differentiation. In both normal tissues and conditions resulting from incorrect or incomplete maturation events, such as metaplastic and neoplastic transformations, TLE expression is elevated and coincides with Notch expression, implicating these molecules in the maintenance of the undifferentiated state in epithelial cells. We also show that TLE1 and TLE2 are organized in a tandem array at chromosomal location 19p13.3, while TLE3 maps to 15q22.

16/7/9 (Item 9 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11799009 BIOSIS Number: 98399009

TAN-1, human homolog of Drosophila Notch, encodes a IKB-like activity
Guan E; Wang J H; Laborda J; Lizzio E; Norcross M; Baeuerle P A; Hoffman T

Cent. Biol. Eval. Res., US FDA, Freiburg, Germany
0 (0). 1995. 693.

Full Journal Title: 9TH INTERNATIONAL CONGRESS OF IMMUNOLOGY. The 9th International Congress of Immunology; Meeting Sponsored by the American Association of Immunologists and the International Union of Immunological

Societies, San Francisco, California, USA, July 23-29, 1995. ix+742p. 9th International Congress of Immunology: San Francisco, California, USA.

ISSN: *****

Language: ENGLISH

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 009 Ref. 161412

16/7/10 (Item 10 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11789761 BIOSIS Number: 98389761

Functional analysis of the TAN-1 gene, a human homolog of Drosophila notch

Aster J; Pear W; Hasserjian R; Erba H; Davi F; Luo B; Scott M; Baltimore D; Sklar J

Div. Molecular Oncol., Dep. Pathol., Brigham and Women's Hosp., Harv. Med. Sch., Boston, MA 02115, USA

0 (0). 1994. 125-136.

Full Journal Title: COLD SPRING HARBOR LABORATORY. Cold Spring Harbor Symposia on Quantitative Biology, Vol. 59. The molecular genetics of cancer; 59th Symposium on Quantitative Biology, Cold Spring Harbor, New York, USA, June 1-8, 1994. xxvii+739p. Cold Spring Harbor Laboratory Press: Plainview, New York, USA. ISBN 0-87969-068-2(paper); ISBN 0-87969-067-4(cloth).

ISSN: 0091-7451

Language: ENGLISH

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 009 Ref. 152164

16/7/11 (Item 11 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11778798 BIOSIS Number: 98378798

Alterations in Notch signaling in neoplastic lesions of the human cervix
Zagouras P; Stifani S; Blaumueller C M; Carcangiu M L; Artavanis-Tsakonas S

Dep. Biol., Yale Univ., New Haven, CT 06536, USA

Proceedings of the National Academy of Sciences of the United States of America 92 (14). 1995. 6414-6418.

Full Journal Title: Proceedings of the National Academy of Sciences of the United States of America

ISSN: 0027-8424

Language: ENGLISH

Print Number: Biological Abstracts Vol. 100 Iss. 005 Ref. 070636

The development of cancer is a cellular process that reflects and is partly driven by alterations in cell determination. Mutations in various molecules responsible for cell determination have been identified as being oncogenic, but little is known about the involvement of normal cell fate-determining mechanisms in the oncogenic process. The Notch pathway defines an evolutionarily conserved, general cell interaction mechanism that controls fundamental aspects of cell determination during vertebrate and invertebrate development. We have explored the involvement of the human Notch pathway in human cervical tissues, which define a cellular environment where cell fate changes take place and where neoplastic conditions have been well characterized. Our evidence suggests that Notch expression is associated with cell populations that are undergoing cell

fate changes and that Notch activity can be used to monitor cell fate abnormalities in cervical as well as other epithelial neoplasias.

16/7/12 (Item 12 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11572298 BIOSIS Number: 98172298

Drosophila in cancer research: The first fifty tumor suppressor genes

Watson K L; Justice R W; Bryant P J

Dep. Molecular Cellular Biol., Harvard Univ., Cambridge, MA 02138, USA

Journal of Cell Science 0 (SUPPL. 18). 1994. 19-33.

Full Journal Title: Journal of Cell Science

ISSN: 0021-9533

Language: ENGLISH

Print Number: Biological Abstracts Vol. 099 Iss. 008 Ref. 112601

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 005 Ref. 069789

In Drosophila, over 50 genes have been identified in which loss-of-function mutations lead to excess cell proliferation in the embryo, in the central nervous system, imaginal discs or hematopoietic organs of the larva, or in the adult gonads. Twenty-two of these genes have been cloned and characterized at the molecular level, and nine of them show clear homology to mammalian genes. Most of these mammalian genes had not been previously implicated in cell proliferation control. Overgrowth in some of the mutants involves conversion to a cell type that, in normal development, shows more cell proliferation than the original cell type. Thus the neurogenic mutants, including Notch, show conversion of epidermal cells to neuroblasts, leading to the 'neurogenic' phenotype of excess nervous tissue. The ovarian tumor mutants show conversion of the female germ line to a cell type resembling the male germ line, which undergoes more proliferation than the female germ line. Mutations of the fat locus cause hyperplastic overgrowth of imaginal discs, in which the epithelial structure is largely intact. The predicted fat protein product is a giant relative of cadherins, supporting indications from human cancer that cadherins play an important role in tumor suppression. Mutations in the lethal(2)giant larvae and lethal(1)discs large genes cause neoplastic overgrowth of imaginal discs as well as the larval brain. The dig gene encodes a membrane-associated guanylate kinase homolog that is localized at septate junctions between epithelial cells. This protein is a member of a family of homologs that also includes two proteins found at mammalian tight junctions (ZO-1 and ZO-2) and a protein found at mammalian synaptic junctions (PSD-95/SAP90). Genes in which mutations cause blood cell overproduction include aberrant immune response-8, which encodes the RpS6 ribosomal protein and hopscotch, which encodes a putative non-receptor protein tyrosine kinase. The gene products identified by ovarian tumor mutants do not show clear amino acid sequence homology to known proteins. Drosophila provides an opportunity to rapidly identify and characterize tumor suppressor genes, many of which have mammalian homologs that might also be involved in cell proliferation control and tumor suppression.

16/7/14 (Item 14 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11492147 BIOSIS Number: 98092147

The human NOTCH1, 2, and 3 genes are located at chromosome positions

9q34, 1p13-p11, and 19p13.2-p13.1 in regions of neoplasia-associated translocation

Larsson C; Lardelli M; White I; Lendahl U

Dep. Cell Mol. Biol., Med. Nobel Inst., Karolinska Hosp., S-171 77 Stockholm, Sweden

Genomics 24 (2). 1994. 253-258.

Full Journal Title: Genomics

ISSN: 0888-7543

Language: ENGLISH

Print Number: Biological Abstracts Vol. 099 Iss. 005 Ref. 062557

In *Drosophila* the Notch gene controls differentiation to various cell fates in many tissues. Three mammalian Notch homologs have recently been identified: Notch 1, 2, and 3. All three homologs are very highly conserved relative to the *Drosophila* Notch gene, which suggests that they are important for cell differentiation in mammals. This notion is supported by the previous finding of a truncated, translocated form of the human NOTCH1 gene (formerly TAN1) in three cases of leukemia. Given this genetic link between NOTCH1 and tumor formation, it is of interest to establish the chromosomal positions of the other two homologs. We report the identification of cosmid clones for the human NOTCH1, 2, and 3 genes. These clones were used as probes in fluorescence in situ hybridization to human metaphase chromosomes, and the results, combined with data from somatic cell hybrid panels, show that the NOTCH2 and 3 genes are located at positions 1p13-p11 and 19p13.2-p13.1, respectively, which are regions of neoplasia-associated translocation.

16/7/15 (Item 15 from file: 5)

DIALOG(R) File 5:BIOSIS PREVIEWS(R)

(c) 1996 BIOSIS. All rts. reserv.

11292950 BIOSIS Number: 97492950

An activated Notch suppresses neurogenesis and myogenesis but not gliogenesis in mammalian cells

Nye J S; Kopan R; Axel R

Dep. Neurol., Inst. Cancer Res., Coll. Physicans Surg., Columbia Univ., New York, NY 10032, USA

Development (Cambridge) 120 (9). 1994. 2421-2430.

Full Journal Title: Development (Cambridge)

ISSN: 0950-1991

Language: ENGLISH

Print Number: Biological Abstracts Vol. 098 Iss. 010 Ref. 128689

P19 cells, a mouse embryonal carcinoma line, can be induced to differentiate into neurons. After induction, however, only a small subpopulation of cells develop as neurons, suggesting that equipotent cells adopt different cell fates. In invertebrate systems, the lin-12-Notch family of genes is thought to control the choice of cell fate. We have therefore asked whether activation of murine Notch (mNotch) regulates neuronal differentiation in P19 cells. We demonstrate that a dominant gain-of-function mutant of mNotch suppresses neurogenesis, as well as myogenesis in P19 cells. Overexpression of the full-length mNotch protein also suppresses neurogenesis. In contrast, the differentiation of glia is not affected by an activated mNotch homologue. These data indicate that mNotch may play a central role in the choice of cell fate in differentiating cells in culture and suggests that mNotch may play a similar role in the choice of fate in the developing mammalian embryo.

16/7/17 (Item 17 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11038694 BIOSIS Number: 97238694

A human homologue of the Drosophila developmental gene, Notch, is expressed in CD34+ hematopoietic precursors

Milner L A; Kopan R; Martin D I K; Bernstein I D

Fred Hutchinson Cancer Res. Cent., 1100 Fairview Ave. N, Seattle WA 98109, USA

Blood 83 (8). 1994. 2057-2062.

Full Journal Title: Blood

ISSN: 0006-4971

Language: ENGLISH

Print Number: Biological Abstracts Vol. 097 Iss. 011 Ref. 155807

Members of the Notch gene family have been shown to mediate cell-fate decisions by multipotent precursors in a number of different systems. To determine whether members of this family might play a similar role in hematopoiesis, we asked if homologues of the Notch gene are expressed in human hematopoietic precursors. Using degenerate oligonucleotides corresponding to conserved amino acid sequences in known Notch homologues as primers for the polymerase chain reaction (PCR), we demonstrated that at least one Notch homologue is expressed in human bone marrow CD34+ cells, a population enriched for hematopoietic precursors. Cloning and sequencing of the PCR products identified this Notch homologue as TAN-1, a member of the Notch family previously cloned from a T-cell leukemia with a translocation involving this gene. Subsequent evaluation of bone marrow hematopoietic cells for TAN-1 expression using a reverse transcription-PCR assay confirmed the expression of TAN-1 in CD34+ hematopoietic precursors, including the immature subset that lacks expression of lineage-associated antigens (CD34+lin-). These findings, together with the known role of Notch homologues in other systems, suggest that members of the Notch family, including TAN-1, may be involved in mediating cell-fate decisions during hemotopoiesis.

16/7/23 (Item 23 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

9139819 BIOSIS Number: 93124819

EXPRESSION OF AN ACTIVATED NOTCH-RELATED INT-3 TRANSGENE INTERFERES WITH CELL DIFFERENTIATION AND INDUCES NEOPLASTIC TRANSFORMATION IN MAMMARY AND SALIVARY GLANDS

JHAPPAN C; GALLAHAN D; STAHL C; CHU E; SMITH G H; MERLINO G; CALLAHAN R
LAB. MOLECULAR BIOL., NATL. CANCER INST., NATL. INST. HEALTH. BETHESDA, MD. 20892, USA.

GENES DEV 6 (3). 1992. 345-355. CODEN: GEDEE

Full Journal Title: Genes & Development

Language: ENGLISH

Expression of the int-3 locus is activated in mouse mammary tumors as a consequence of insertional mutagenesis by the mouse mammary tumor virus (MMTV). Integration of the MMTV provirus into the into-3 locus promotes the transcription and translation of flanking cellular int-3 sequences sharing significant homology with the intracellular domain of the nuerogenic Notch gene of Drosophila, and with the yeast cell cycle regulatory genes cdc10 and SW16. To determine the in vivo consequences of activated int-3 expression, transgenic mice were generated harboring a genomic tumor DNA

fragment consisting of the MMTV LTR and the flanking cellular int-3 sequences. All six int-3 founder transgenic mice and the progeny of one established line exhibited similar dramatic phenotypic abnormalities in tissues in which the transgene was expressed. Focal and often multiple poorly differentiated mammary and salivary adenocarcinomas appeared in the majority of transgenic mice between 2 and 7 months of age. Significantly, mammary glands were arrested in development and were lactation deficient in all female int-3 mice. The salivary glands, glands of the nasal mucosa and maxillary sinus, the extraorbital lacrimal glands, and the Harderian glands of juvenile and adult transgenic mice all contained proliferating immature ductule cells and were incompletely differentiated. In addition, all male int-3 transgenic mice were sterile, apparently the result of severe hyperplasia of the epididymis. These findings demonstrate in vivo that expression of the activated Notch-related int-3 gene causes deregulation of normal developmental controls and hyperproliferation of glandular epithelia.

16/7/25 (Item 25 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

8646875 BIOSIS Number: 92111875
TAN-1 THE HUMAN HOMOLOG OF THE DROSOPHILA NOTCH GENE IS BROKEN BY
CHROMOSOMAL TRANSLOCATIONS IN T LYMPHOBLASTIC NEOPLASMS
ELLISEN L W; BIRD J; WEST D C; SORENG A L; REYNOLDS T C; SMITH S D; SKLAR
J

STANFORD UNIV. SCH. MED., STANFORD, CALIF. 94305.

CELL 66 (4). 1991. 649-662. CODEN: CELLB

Full Journal Title: Cell

Language: ENGLISH

Previously we described joining of DNA in the .beta. T cell receptor gene to DNA of an uncharacterized locus in a t(7;9)(q34;q34.3) chromosomal translocation from a case of human lymphoblastic leukemia (T-ALL). We now show that the locus on chromosome 9 contains a gene highly homologous to the Drosophila gene Notch. Transcripts of the human gene, for which we propose the name TAN-1, and its murine counterpart are present in many normal human fetal and adult mouse tissues, but are most abundant in lymphoid tissues. In t(7;9)(q34;q34.3) translocations from three cases of T-ALL, the breakpoints occur within 100 bp of an intron in TAN-1, resulting in truncation of TAN-1 transcripts. These observations suggest that TAN-1 may be important for normal lymphocyte function and that alteration of TAN-1 may play a role in the pathogenesis of some T cell neoplasms.

16/7/40 (Item 1 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1996 Elsevier Science B.V. All rts. reserv.

9978677 EMBASE No: 96166337

T cell leukemia-associated human Notch/translocation-associated Notch homologue has IkappaB-like activity and physically interacts with nuclear factor-kappaB proteins in T cells

Guan E.; Wang J.; Laborda J.; Norcross M.; Baeuerle P.A.; Hoffman T.

Division of Monoclonal Antibodies, Federal Drug Administration, Ctr. for Biologics Evaluation/Res., 8800 Rockville Pike, Bethesda, MD 20892 USA

Journal of Experimental Medicine (USA) , 1996, 183/5 (2025-2032) CODEN: JEMEA
ISSN: 0022-1007

LANGUAGES: English SUMMARY LANGUAGES: English

Translocation-associated Notch homologue (TAN-1), a gene originally cloned from the translocation breakpoint of a human T cell leukemia carrying a 9:7(q34.3) translocation, encodes a protein belonging to the Notch/Lin-12/Glp-1 receptor family. These receptors mediate the specification of numerous cell fates during development in invertebrates and vertebrates. The intracellular portion of Notch/TAN-1 contains six ankyrin repeats that are similar to those found in cytoplasmic IkappaB proteins. IkappaB proteins are specific inhibitors of nuclear factor (NF)-kappaB/Rel transcription factors. Here we show that TAN-1 has functional properties of an IkappaB-like regulator with specificity for the NF-kappaB p50 subunit. A recombinant polypeptide corresponding to the cytoplasmic portion of TAN-1 (TAN-1(C)) specifically inhibited the DNA binding of p50-containing NF-kappaB complexes. When overexpressed in an appropriate cell line, TAN-1(C) prevented kappaB-dependent transactivation in transient reporter gene assays in a fashion similar to the structurally related protein, Bcl-3. TAN-1(C) could activate kappaB-dependent gene expression by attenuating the inhibitory effect of an excess of p50 homodimers. Immunoprecipitation experiments showed that the TAN-1 from a T cell line is associated with NF-kappaB containing p50 and p65 subunits. These observations indicate that TAN-1(C) may directly engage NF-kappaB transcription factors and modulate nuclear gene expression.

?LOGOFF